REMARKS

After entry of this amendment, claims 2-15, 23, 28, 31-33, and 50-65 are pending. The claims have been amended without prejudice or disclaimer to correct the dependency in light of the cancellation of claim 1 made in the Response dated August 13, 2009. Support is found *inter alia* in the original claims. No new matter has been added.

Claim Rejections – 35 U.S.C. § 112

Claims 2-15, 23, 28, 31-33 and 50-65 are rejected under 35 U.S.C. § 112, first paragraph, for allegedly lacking an enabling disclosure.

The Examiner asserts that deregulating including overexpressing any of the genes encoding enzymes and/or proteins of the pantothenate biosynthetic pathway and overexpressing any of the genes encoding enzymes and/or proteins of the methylenetetrahydrofolate (MTF) biosynthetic pathway will involve manipulating any of the genes encoding any and all enzymes and/or proteins associated with any of the entire metabolism pathways. To do so, the Examiner contends that one skilled artisan would need to have the knowledge of the pantothenate biosynthetic genes, the isoleucine-valine (ilv) biosynthetic genes, and the MTF biosynthetic genes. The Examiner, however, alleges that many of these genes have not been discovered. The Examiner further contends that one skilled artisan would also need to have the knowledge of genes regulating pantothenate kinase or manipulating the kinase activity. Asserting that mere knowledge of the biochemical or metabolic pathways leading to the production of a compound is not sufficient to overexpress any of the genes associated with the pathways, the Examiner concludes that undue experimentation would be required to search and screen for any deregulation of any enzymes and/or proteins of the above pathways and to determine whether such deregulation leads to an enhanced production of pantothenate. Applicants strongly disagree and traverse the rejection for the reasons already of record and for the following additional reasons.

It is noted initially that the disclosure provided in the specification is presumptively enabling. The manner of making and using the claimed invention must be taken as in compliance with the first paragraph of 35 U.S.C. §112, unless there is objective evidence or scientifically based reasoning inconsistent with the specification. *See In re Marzocchii and Horton*, 169 U.S.P.Q. 367 (C.C.P.A. 1971). "It is the Patent Office's burden to present evidence

that there is some reason to dispute the enablement provided in the specification. Unsupported speculation or conjecture on that the invention 'might not work' will not support a rejection based on 35 U.S.C. §112, first paragraph." *Id.* Simply pointing to the absence of a working example provides neither objective evidence nor reasoning in support of the rejection, and accordingly, a *prima facie* case of non-enablement on this ground has not been made out. Moreover, there has never been a requirement that every species encompassed by a claim must be disclosed or exemplified in a working example. *In re Angstadt*, 537 F.2d 498 (CCPA 1976). Additionally, even though practicing the full scope of the claims might have required some amount of experimentation, if the experimental techniques are well-known in the art, the experimentation is routine and not undue. See *Ex parte Kubin*, 83 USPQ2d 1410 (B.P.A.I. 2007), *aff'd on other ground*, 90 USPQ 2d 1417 (Fed. Cir. 2009).

Here, the claims are directed to a process for the enhanced production of pantothenate comprising culturing a recombinant microorganism of certain genera having the characteristics as specified in the claims, i.e. a deregulated pantothenate biosynthetic pathway and a deregulated MTF biosynthetic pathway, under conditions such that pantothenate production is enhanced as compared to a corresponding wild-type microorganism. The claims further specify that the microorganism suitable for practicing the claimed process should be derived from the genus of *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Streptomyces*, *Salmonella*, *Escherichia*, *Klebsiella*, *Serratia*, *Proteus*, or *Saccharomyces*. Additionally, the "deregulated MTF biosynthetic pathway" is further defined as being by way of overexpressing at least one of the *gcv*, *serA*, *serC*, *serB*, *glyA*, *sul*, *fol*, *mtrA*, *pab*, *panB* and *purR* genes derived from *Bacillus*, *Corynebacterium*, *Lactobacillus*, *Lactococci*, or *Streptomyces*. Thus, the claims are not unduly broad as alleged by the Examiner. Rather, the subject matter encompassed by the present claims are well defined and the specification, together with the knowledge of the art, provides sufficient guidance, including working examples, to make and use the claimed process without undue experimentation.

It is noted initially that, as acknowledged by the Examiner, the specification provides working examples demonstrating how to make and use the claimed process by overexpressing several different genes in *Bacillus*. The working examples provided in the specification describe, in detail, how to generate *Bacillus* strains with a deregulated pantothenate biosynthetic pathway (Example I), how to isolate genes involving in the MTF biosynthetic pathway

(Examples III and IV) or regulating the MTF production (Example VI), how to construct an expression cassette to overexpress a gene in a microorganism (e.g. Examples III and IV), how to produce *Bacillus* strains with a deregulated MTF biosynthetic pathway (Examples III, IV, and VI-IX), and how to detect the effect of increased gene expression on pantothenate production (Examples I and III-V). The methodologies described in the specification are applicable not only to the genes exemplified in the working examples, but also any other genes that are identified as being involving in the pantothenate biosynthetic pathway and/or the MTF biosynthetic pathway. Moreover, the specification further provides guidance as to elements necessary to express genes in a microorganism such as promoter sequences (page 20, lines 1-27), selectable marker sequences (page 22, lines 10-24), and artificial ribosome binding sites (page 21, lines 2-31). Additionally, Applicants submit that constructing vectors for overexpressing genes in microorganisms such as those recited in the claims is routinely used by one skilled in the art. Thus, although some testing and screening would be required to identify a recombinant microorganism useful for practicing the claimed process, such testing and screening would not be extensive and is routine in nature, and thus, not undue. *See Ex parte Jackson*, 217 USPQ 804, 807 (1982) (the test for "undue experimentation" is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine); *see also Ex parte Kubin*, 83 USPQ2d 1410 (B.P.A.I. 2007), *aff'd on other ground*, 90 USPQ 2d 1417 (Fed. Cir. 2009) (even though practicing the full scope of the claims might have required some amount of experimentation, if the experimental techniques are well-known in the art, the experimentation is routine and not undue).

Moreover, it is further noted that, as well known in the art, pantothenate is naturally synthesized in bacteria, fungi, and plants. Thus, the microorganisms recited in the claims as being suitable for practicing the claimed process are pantothenate-producing microorganisms. As such, these microorganisms will naturally possess a pantothenate biosynthetic pathway as well as genes required for the production of pantothenate. For the same reason, these microorganisms would also be expected to naturally possess a MTF biosynthetic pathway as well as genes involved in this pathway because, as described in the specification, the MTF biosynthetic pathway supplies a substrate for the production of pantothenate. As demonstrated in Figures 1 and 2 of the specification, the pantothenate biosynthetic pathway, the isoleucine-valine (ilv) biosynthetic pathway, and the MTF biosynthetic pathway, including enzymes and/or

proteins involved in each step leading to the production of pantothenate, are well known in the art. This is further evidenced by Sahm *et al.* (Applied and Environmental Microbiology, 1999, 65(5): 1973-1979, cited in IDS dated October 3, 2005; see Fig. 1 at page 1974) and Voet *et al.* (Biochemistry, Second Edition, 1995, John Wiley & Sons, Inc., pp. 761-764, cited in the Response dated August 13, 2009; see particularly Figure 24-39 at page 763). Thus, the specification and the art provide the knowledge of the pantothenate biosynthetic genes, the isoleucine-valine (ilv) biosynthetic genes, and the MTF biosynthetic genes necessary for one skilled artisan to practice the claimed process.

Additionally, Applicants also note that, contrary to the Examiner's assertion, many of the genes involved in the pantothenate biosynthetic pathway and the MTF biosynthetic pathway, including the genes recited in the claims for deregulating the MTF pathway, have been isolated from various organisms. For instance, the *serA* gene has been isolated from species such as *E. coli*, *Bacillus*, and *Saccharomyces cerevisiae*. Similarly, the *glyA* gene has been isolated from species such as *Saccharomyces cerevisiae*, *E. coli*, *Salmonella typhimurium*, and *Streptomyces coelicolor*. A Table providing examples of genes recited in the claims that have been isolated is attached herewith for the Examiner's reference. The articles or references cited in the enclosed Table can be made available for the Examiner's review upon request. As a person of ordinary skill in the art involved in the production of pantothenate would have an advanced degree (likely a Ph.D. degree) in the molecular biological science and several years of work experience, identifying functional homologues or equivalents from other microorganisms would be simply routine practice. For example, degenerated primers may be used to isolate homologues or equivalents of a known gene from the gene pool of another microorganism, which can then be tested for its enzymatic function in a complementation assay. Methodologies or techniques required for identifying and isolating homologues or equivalents of a known gene are well known in the art and within the knowledge and experience level of one of ordinary skill. Moreover, with the increasing availability of complete genome sequences of many other microorganisms, one skilled in the art would be able to easily identify and screen for equivalent genes suitable for practicing the claimed process.[1] With this regard, Applicants note that detailed

[1] According to the NCBI website, more than 1000 prokaryotic genomes are now completed and available in the Genome database, including the genera of microorganisms recited in the claims, i.e. *Bacillus* (106 entries), *Corynebacterium* (25 entries), *Lactobacillus* (59 entries), *Lactococci* (4 entries), and *Streptomyces* (22 entries). See http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html.

guidance as to how to isolate genes from other microorganisms is provided throughout the specification, including working examples, as discussed above. See e.g., Specification at page 17, lines 4-15, and Examples III and IV. Accordingly, only very extensive and very unpredictable experimentation would rise to the level of "undue" in this art, which is clearly not present in the instant application.

The routine nature of practicing the claimed subject matter based on the detailed guidance provided in the specification and the knowledge of the art is further evidenced by, for example, Collet *et al.* (FEBS Letters, 1997, 408: 281-284, copy attached; hereinafter "Collet") and Ho *et al.* (Plant J., 1998, 16(4): 443-452, copy attached; hereinafter "Ho"). In Collet, a human *serB* gene was cloned by PCR amplification using primers derived from human EST sequences which were identified in a Blast search using the *serB* sequence from *S. mansoni*. The enzyme encoded by the isolated human *serB* gene shares 30% and 40% sequence identity with the enzymes from *E. coli* and *S. mansoni*, respectively, and was shown to have the phosphoserine phosphatase activity (Collet at page 283, left Col.). Similarly, in Ho, an *Arabidopsis serC* gene was isolated from a cDNA library using the insert of an *Arabidopsis* EST clone showing 55-67% homology to bacterial and yeast *serC* protein as probe. The *Arabidopsis serC* protein shares 34-71% homology with bacterial, yeast and spinach *serC* proteins (Ho at page 444, right Col., 2nd full paragraph) and was shown to have the phosphoserine aminotransferase activity (Ho at page 446, right Col., 2nd full paragraph). These results, although obtained from organisms different from those recited in the claims, demonstrate nevertheless that it is merely routine for one skilled in the art to identify and isolate a gene involved in the pantothenate biosynthetic pathway and/or the MTF biosynthetic pathway.

Applicants additionally note that, in both Collet and Ho, sequence alignments between the *serB* proteins and the *serC* proteins, respectively, from various organisms were provided. For example, in Collet, the *serB* proteins from rat, *S. mansoni*, human, *E. coli*, *H. influenzae*, *Methanococcus jannaschii* and yeast were aligned, and conserved regions were identified (Collet at page 282, Figure 1). Likewise, in Ho, the *serC* proteins from *Arabidopsis*, spinach, *Bacillus circulans*, *E. coli*, rabbit, and yeast were aligned, and signature sequences for *serC* protein as well as conserved sequence motifs were identified (Ho at page 445, Figure 2). Similar disclosure and identification of conserved sequences were also available for other genes involved in the pantothenate biosynthetic pathway and/or the MTF biosynthetic pathway, for example, *glyA*

gene (see e.g., Vatcher et al., J. of Biological Chemistry, 1998, 273(11): 6066-6073, copy attached) and *serA* gene (see e.g., Achouri *et al.*, Biochem. J., 1997, 323: 365-370, copy attached). Based on the abundant information available in the art as illustrated in the above references, one skilled artisan would be able to easily identify the consensus sequence(s) for the gene involved in the pantothenate biosynthetic pathway and/or the MTF biosynthetic pathway, including the gene recited in the claims for deregulating the MTF pathway, and to clone such a gene using the methodologies or techniques taught in the specification or known in the art. As demonstrated by Collet and Ho, the testing and screening required for such identification and isolation is simply routine, and not undue.

In view of the detailed description, guidance and working examples provided in the specification, as well as the state and knowledge of the art and high level of skill, Applicants respectfully submit that the specification enables the full scope of the claims without undue experimentation as further evidenced by the aforementioned references. On these facts, a proper analysis of the relevant factors supports enablement. *In re Wands*, 858 F.2d 731, 737 (Fed. Cir. 1988). For at least the above reasons and for the reasons already of record, reconsideration and withdrawal of the enablement rejection is respectfully requested.

Moreover, Applicants note again that the initial burden is on the Patent Office to present objective evidence or scientifically based reasoning to dispute the presumptively enabling disclosure provided in the specification in light of the knowledge of the art. Simply pointing to the absence of a working example is neither objective evidence nor reasoning to support a nonenablement rejection. Unsupported speculation or conjecture on that the invention "might not work" will also not support such a rejection. Additionally, some amount of experimentation will not defeat enablement so long as it is routine and not undue. Since the Examiner has not provided objective evidence or scientifically based reasoning to dispute the enablement provided in the specification and the art, a *prima facie* case of non-enablement has not been established. For this additional reason, the nonenablement rejection should be withdrawn.

Claim Rejections – Double Patenting

Claims 2-15, 23, 28, 31-33 and 50-65 are rejected on the ground of nonstatutory obviousness-type double patenting over claims 1-17 of U.S. Pat. No. 7,291,789 and claims 1-34 of U.S. Pat. No. 7,244,593. Claims 2-15, 23, 28, 31-33 and 50-65 are further provisionally

rejected for obviousness-type double patenting over claims 2-6, 11-26, 29-32 and 35-40 of co-pending Application No. 11/879,143. Applicants will consider filling a terminal disclaimer upon an indication that the claims are allowable.

Power of Attorney

A Transmittal of Power of Attorney with Change of Correspondence Address together with the executed Power Of Attorney To Prosecute Applications Before the USPTO and a Statement Under 37 CFR § 3.73(b) were filed in the present application on March 12, 2009. Applicants respectfully request that all pertinent U.S. Patent and Trademark Office records relating to the subject application be changed accordingly.

CONCLUSION

In view of the above remarks and further in view of the above amendments, Applicants respectfully request withdrawal of the rejection and allowance of the claims. If any outstanding issues remain, the Examiner is invited to telephone the undersigned at the number given below.

This response is filed within the three-month period for response from the mailing of the Office Communication, to and including March 1, 2010. No fee is believed due. However, if a fee is due, please charge our Deposit Account No. 03-2775, under Order No. 13311-00036-US from which the undersigned is authorized to draw.

Respectfully submitted,

By /Hui-Ju Wu/

Hui-Ju Wu, Ph.D.
Registration No.: 57,209
CONNOLLY BOVE LODGE & HUTZ LLP
1007 North Orange Street; P. O. Box 2207
Wilmington, Delaware 19899-2207
(302) 658-9141
(302) 658-5614 (Fax)
Attorney for Applicants

#741506

Attachments:

1. Table exemplifying genes recited in the claims that have been isolated.
2. Collet *et al.*, FEBS Letters, 1997, 408: 281-284.
3. Ho *et al.*, Plant J., 1998, 16(4): 443-452.
4. Vatcher *et al.*, J. of Biological Chemistry, 1998, 273(11): 6066-6073.
5. Achouri *et al.*, Biochem. J., 1997, 323: 365-370.

Attachment 1:

Gene	Organism	Reference(s)
serA	*E. coli*	Tobey *et al.*, J. Biol. Chem., 1986, 261: 12179-12183
	H. influenzae	Genbank accession No. L45106
	Yeast	Genbank accession No. P40054
	Bacillus	Genbank accession No. L47648
	Rat	Achouri *et al.*, Biochem. J., 1997, 323: 365-370
serB	*E. coli*	Genbank accession No. P06862
	Bacillus subtilis	Genbank accession No. NP_391291
	H. influenzae	Genbank accession No. L42023
	Human	Collet *et al.*, FEBS Letters, 1997, 408: 281-284
serC	*H. influenzae*	Genbank accession No. P44336
	Arabidopsis	Ho *et al.*, Plant J., 1998, 16(4): 443-452
	Spinach	Genbank accession No. D84061
	Bacillus circulans	Genbank accession No. Z46432
gcv	*E. coli*	Okamura-Ikeda *et al.*, Eur. J. Biochem., 1993, 216: 539-548
glyA	*E. coli*	Genbank accession No. AAC75604
	C. elegans	Vatcher *et al.*, J. Biol. Chem., 1998, 278(11): 6066-6073
	Salmonella typhimurium	Steiert *et al.*, DNA Seq., 1990, 1(2): 107-113
	Streptomyces coelicolor	Smith *et al.*, Mol. Gen. Genet., 1988, 211(1): 129-137
sul/sfi	*E. coli*	Jones *et al.*, PNAS, 1985, 82: 6045-6049 Kanemori *et al.*, J. Bacteriol., 1999, 181(12): 3674-3680
fol	*E. coli*	Genbank accession No. NP_414590
	Mycobacteria	Barrow *et al.*, U.S. Pat. No. 6,229,001
mtrA	*Methanobacterium thermoautotrophicum*	Stupperich *et al.*, Eur. J. Biochem., 1993, 217(1): 115-121
	Methanopyrus kandleri	Harms *et al.*, Eur. J. Biochem., 1997, 250: 783-788
	Methanosarcina barkeri	Harms *et al.*, Eur. J. Biochem., 1997, 250: 783-788
pab	Yeast	Genbank accession No. P37254
panB	*E. coli*	Jones *et al.*, J. Bacteriol., 1993, 175(7): 2125-2130
purR	*E. coli*	He *et al.*, J. Bacteriol., 1992, 174(1): 130-136

Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate

Jean-François Collet[a], Isabelle Gerin[a], Mark H. Rider[b], Maria Veiga-da-Cunha[a], Emile Van Schaftingen[a,*]

[a]*Laboratory of Physiological Chemistry, ICP and University of Louvain, B-1200 Brussels, Belgium*
[b]*Hormone and Metabolic Research Unit, ICP and University of Louvain, B-1200 Brussels, Belgium*

Received 15 January 1997; revised version received 6 April 1997

Abstract **We report the sequence of the cDNA encoding human L-3-phosphoserine phosphatase. The encoded polypeptide contains 225 residues and shows 30% sequence identity with the *Escherichia coli* enzyme. The human protein was expressed in a bacterial expression system and purified. Similar to known L-3-phosphoserine phosphatases, it catalyzed the Mg^{2+}-dependent hydrolysis of L-phosphoserine and an exchange reaction between L-serine and L-phosphoserine. In addition we found that the enzyme was phosphorylated upon incubation with L-[^{32}P]phosphoserine, which indicates that the reaction mechanism proceeds via the formation of a phosphoryl-enzyme intermediate. The sensitivity of the phosphoryl-enzyme to alkali and to hydroxylamine suggests that an aspartyl- or a glutamyl-phosphate was formed. The nucleotide sequence of the cDNA described in this article has been deposited in the EMBL data base under accession number Y10275.**

1. Introduction

L-3-Phosphoserine phosphatase is the enzyme responsible for the third and last step in L-serine formation. This enzyme not only catalyzes the Mg^{2+}-dependent hydrolysis of L-phosphoserine but also an exchange reaction between L-serine and L-phosphoserine [1,2], suggesting that its reaction mechanism proceeds through the formation of a phosphoryl-enzyme intermediate; the latter has, however, never been directly demonstrated. As L-3-phosphoserine phosphatase has been reported to be deficient in one case of L-serine deficiency [3], it was of interest to determine the sequence of the human enzyme. Databanks contain DNA sequences of several bacteria and of two eukaryotes (*Saccharomyces cerevisiae* and *Schistosoma mansoni*) that are homologous to *SerB*, the gene encoding *Escherichia coli* L-3-phosphoserine phosphatase [4].

In this paper we report the sequence of the human enzyme. We have also expressed an active protein and show that it forms a phosphoenzyme when incubated with its substrate.

2. Material and methods

2.1. Materials

L-[^{32}P]Phosphoserine was synthesized by incubating 1 mM L-serine with 10 μM ^{32}P-labeled inorganic pyrophosphate (0.10 mCi), 2 mM magnesium acetate, 1 mM dithiothreitol, 25 mM Tris-HCl (pH 7.1), and 0.03 U pyrophosphate: L-serine phosphotransferase [5] for 40 min at 30°C in a final volume of 3 ml; it was purified as in [6]. Radioactive compounds and Thermosequenase were from Amersham and *Pwo* polymerase from Boehringer. Chemicals were from Sigma or Merck. Clones R14208, N23530, N42133, W05752, H38879 and T82144 from the IMAGE consortium [7], were kindly provided by the UK HGMP Resource Centre.

2.2. Amplification of cDNA

DNA obtained from human cDNA libraries (from renal cell carcinoma line LE9211-RCC [8], from the urinary bladder transitional-cell carcinoma LB831-BLC, or from EBV-transformed lymphoblastoid cell line LG2-EBV [9]) was amplified with *Pwo* polymerase (a polymerase with proofreading activity) using two primers chosen as described in Section 3. The first (GTGCATATGGTCTCCCACTCAGAGCTG) had a start codon (underlined) inserted in a *Nde*I site and the second (ACGGATCCTCATCTGAAGTTGTTTGGAGC) corresponded to nucleotides 882–902 of the sequence that we report in the EMBL databank, flanked by a *Bam*HI site. The PCR-amplified product obtained with the human kidney library was purified by agarose gel electrophoresis, inserted in the *Eco*RV site of pBluescript and sequenced.

2.3. Expression and purification of recombinant human L-3-phosphoserine phosphatase

The insert of the resulting plasmid was excised with *Nde*I and *Bam*HI restriction enzymes and ligated in the expression vector pET3a [10]. BL21(DE3)pLysS cells were transformed with this plasmid and were aerobically grown in M9 medium at 37°C until A_{600} reached 0.5–0.6. The culture was then maintained on ice for 15 min, after which time isopropylthiogalactoside was added to a final concentration of 0.4 mM. After 20 h of incubation at 37°C, bacterial extracts were prepared as described [11], the lysis buffer being supplemented with 5 μg/ml of antipain and 5 μg/ml of leupeptin. The extract was centrifuged for 40 min (20 000×*g* at 4°C). The resulting supernatant (25 ml) was diluted 3-fold with buffer A (10 mM Tris-HCl (pH 8.5), 1 mM dithiothreitol, 1 μg/ml leupeptin, 1 μg/ml antipain) and applied onto a DEAE-Sepharose column (1.6×10 cm). The column was washed with 100 ml of buffer A and protein was eluted with a NaCl gradient (0–400 mM in 250 ml of buffer A). L-3-Phosphoserine phosphatase came out at a Na^+ concentration of 250 mM. The active fractions were pooled and concentrated 4-fold by ultrafiltration in an Amicon cell (YM-10 membrane). Two milliliters of this preparation were further purified by gel filtration on Sephacryl S-200 (1.6×50 cm) in buffer B (20 mM Hepes (pH 7.5), 1 mM dithiothreitol, 100 mM KCl, 0.5 μg/ml leupeptin and 0.5 μg/ml antipain) at a flow rate of 0.25 ml/min.

2.4. Purification and sequencing of rat liver L-3-phosphoserine phosphatase

L-3-Phosphoserine phosphatase was purified from rat liver by a previously described procedure [12], followed by a gel filtration on Sephacryl S-200 equilibrated with buffer B. The active fractions were submitted to SDS-PAGE. Staining with Coomassie Blue revealed the presence of about 10 bands with M_r 20 000–70 000. The gel was sliced, and the polypeptides were extracted and subjected to a renaturation procedure [13]. Phosphoserine phosphatase activity was observed only in the fraction corresponding to a 25 000 M_r polypeptide. Tryptic peptides were obtained from this protein and sequenced as described [14].

*Corresponding author. UCL 7539, Avenue Hippocrate 75, B-1200 Brussels, Belgium. Fax: (32) 2-7647598.
E-mail: vanschaftingen@bchm.ucl.ac.be

PII S0014-5793(97)00438-9

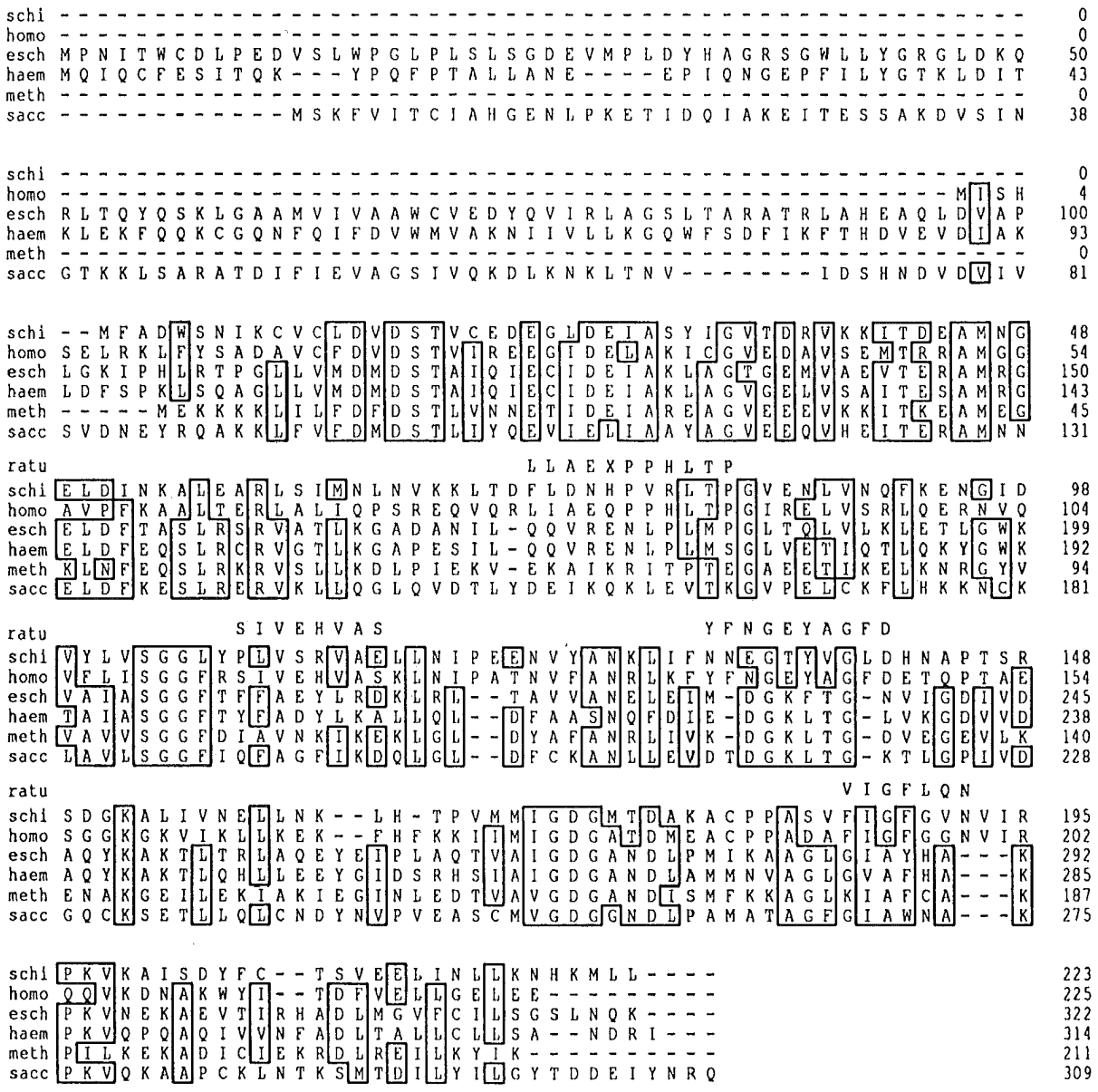


Fig. 1. Alignment of l-3-phosphoserine phosphatase sequences from various species (ratu: *Rattus norvegicus*; schi: *Schistosoma mansoni*; homo: *Homo sapiens*; esch: *Escherichia coli*; haem: *Haemophilus influenzae*; meth: *Methanococcus jannaschii*; sacc: *Saccharomyces cerevisiae*). The tryptic peptides are shown for the rat enzyme. Similar and conserved residues are shown in boxes.

2.5. Enzyme assays

In the course of its purification, L-3-phosphoserine phosphatase was assayed at 30°C by the release of Pi from unlabelled L-phosphoserine in an assay mixture (250 μl) containing 25 mM Mes (pH 6.5), 5 mM $MgCl_2$, 1 mM dithiothreitol, 5 mM L-phosphoserine, 0.1 mg/ml bovine serum albumin and 1–10 mU L-3-phosphoserine phosphatase. Reactions were stopped by the addition of 250 μl of 10% trichloroacetic acid (TCA) and Pi was measured [15]. One unit of enzyme is the amount that catalyzes the conversion of 1 μmol of substrate per minute under these conditions. For kinetic studies, the enzyme was assayed by the release of [^{32}P]Pi from L-[^{32}P]phosphoserine [16] in the assay mixture described above. The exchange reaction was measured by the incorporation of L-[^{14}C]serine into L-phosphoserine in an assay mixture containing 25 mM Hepes (pH 7.5), 1 mM dithiothreitol, 1 mM $MgCl_2$, 1 mM L-phosphoserine, 0.5 mM L-serine and 20 000 cpm L-[^{14}C]serine in a final volume of 100 μl. The incubation was arrested by spotting 25 μl of the reaction mixture on Whatman DE-81 papers, which were washed in water, dried and counted for radioactivity in the presence of 5 ml of HiSafe 2 scintillant.

2.6. Demonstration of the formation of the phosphoenzyme

Unless otherwise indicated, about 0.15 U of recombinant L-3-phosphoserine phosphatase was incubated at 0°C in a mixture containing 50 mM Hepes (pH 7.5), 1 mM dithiothreitol, 1 μM L-phosphoserine, 60 000 cpm of L-[^{32}P]phosphoserine, 5 mM $MgCl_2$, 20 mM L-serine and 1 mg/ml bovine serum albumin in a volume of 100 μl. The reaction was stopped at the indicated times by the addition of 250 μl of 5% TCA and filtration on a polyethersulfone membrane (Supor® 200, from Gelman). The membrane was washed with 10 ml of TCA 5% and counted for radioactivity with 5 ml of HiSafe 2 scintillant.

2.7. Other methods

Protein was measured according to Bradford [17] with bovine gamma-globulin as a standard. Sequencing was performed using T7 Ther-

mo Sequenase (Amersham), fluorescent primers and the LI-COR automated DNA sequencer 4000L.

3. Results

3.1. Identification and sequencing of human cDNAs

The predicted amino acid sequence of the *S. mansoni* enzyme, most likely the closest to that of the human enzyme, was used to perform a Blast search [18]. Several human ESTs (expressed-sequence tags) were identified, which corresponded to the 5′ (IMAGE Consortium [LLNL] cDNA clones R14208; N23530; N42133; W05752) or the 3′ end (clones H38879 and T82144) of the parasite enzyme's open reading frame. Oligonucleotide primers derived from these sequences were used in PCR reactions to amplify human cDNAs derived from bladder or kidney tumor cells, or of lymphocytes. In all cases one single fragment of 725 bp was obtained. The fragment obtained by amplification of the kidney cell cDNA was cloned in pBluescript and sequenced. Several of the above-mentioned clones were also sequenced. The composite sequence of the reconstituted human cDNA is reported in the EMBL databank. Note that clones N42133 and W05752 lacked nucleotides 281 to 757, and that in clone N23530, there was a 122 bp insert homologous to the *Alu* I family between nucleotides 280 and 758.

The sequence totals 1600 bp; the 5′ and 3′ non-coding regions amount to 187 and 735 bp, respectively. The ATG codon opens a reading frame of 675 bp, encoding a protein of 225 amino acids. This protein shows 30% and 40% sequence identity with the enzymes from *E. coli* and *S. mansoni*, respectively. It shows also a high degree of identity with several peptides derived from L-3-phosphoserine phosphatase purified from rat liver (Fig. 1).

3.2. Expression and characterization of a recombinant protein

To check that the sequence encoded L-3-phosphoserine phosphatase, the open reading frame was inserted in pET3a [10] and expressed in *E. coli*. Induction of the cells with isopropylthiogalactoside led to the appearance of a 25000 M_r polypeptide which after 20 h represented about 5% of the soluble proteins. L-3-Phosphoserine phosphatase was purified by chromatography on DEAE-Sepharose and Sephacryl S-200, from which it was eluted with an apparent molecular mass of 50000 Mr, indicating a dimeric structure (not shown). The purified protein had a specific activity of 6 U/mg protein, the yield of the purification being 25%.

The purified protein was dependent for its activity on the presence of Mg^{2+}, which half-maximally stimulated the enzyme at 5 mM. It displayed a K_m of 20 μM for L-phosphoserine in the presence of 1 mM Mg^{2+}, was non-competitively inhibited by L-serine (K_i = 0.5 mM) and catalyzed an exchange reaction (incorporation of L-[^{14}C]serine into L-phosphoserine) at a rate about 50% that of the hydrolysis (3 μmol/min/mg protein). These properties are similar to those of other L-3-phosphoserine phosphatases [1,2,19]. Other phosphate esters (glucose 6-phosphate, fructose 6-phosphate, glycerol 2- or 3-phosphate and 3-phosphoglycerate, all tested at 5 mM) were not significantly hydrolyzed by the purified preparation.

3.3. Demonstration of the formation of a phosphoenzyme

Upon incubation of human L-3-phosphoserine phosphatase with L-[^{32}P]phosphoserine, a phosphoenzyme was formed.

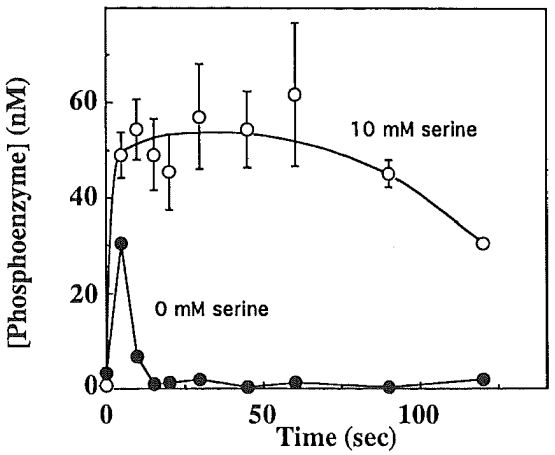


Fig. 2. Time-course of the labelling of the rat liver enzyme with L-[^{32}P]phosphoserine. L-3-Phosphoserine phosphatase (0.15 U) was incubated at 0°C in the presence of radiolabelled L-phosphoserine with (○) or without (●) 10 mM L-serine.

This could be demonstrated using a method in which protein was precipitated in acid and immediately isolated by filtration. Other methods in which filters are washed for prolonged time (30–60 min) [20,21] yielded insignificant incorporation, presumably because of the lability of the phosphoenzyme (see below). As shown in Fig. 2, the formation of the phosphoenzyme was greatly favored by L-serine which acted most likely by slowing down the degradation of L-phosphoserine. In the presence of 10 mM L-serine, the K_m for the formation of this phosphoenzyme was 5 μM and a maximum of 2.5 nmol of phosphate was incorporated per milligram of protein, indicating a stoichiometry of 0.06 mole per mole of enzyme subunit (not shown). SDS-PAGE [22] under denaturing conditions (at 0°C, to minimize hydrolysis) confirmed that the phosphate was only incorporated into the 25000 M_r polypeptide (not shown). The phosphoenzyme was acid labile, being 50% hydrolyzed after 50 min of incubation in 5% TCA at 50°C. It was extremely labile in alkali, since it was completely lost when resuspended in 1 M NaOH at 0°C and reprecipitated in TCA. It was also completely hydrolyzed when resuspended in 0.2 M NH_2OH at pH 5.3 (not shown) and maintained for 10 min at 20°C.

4. Discussion

We report here the identification of clones encoding human L-3-phosphoserine phosphatase. This identification rests on the similarity of the encoded protein with the *E. coli* enzyme as well as with tryptic peptides derived from the purified rat liver enzyme. The identity was confirmed by expression of the cDNA and demonstration that the encoded protein behaved as a specific L-3-phosphoserine phosphatase with properties similar to those of the enzyme present in mammalian tissues.

The fact that L-3-phosphoserine phosphatase catalyzes an exchange reaction between L-serine and L-phosphoserine and that it is non-competitively inhibited by L-serine suggested that the mechanism of this enzyme involved the formation of a phosphoenzyme. This is also consistent with its sensitivity to vanadate [23,24]. We show in this work that, indeed, a

phosphorylated enzyme is formed upon incubation with the substrate.

Many phosphotransferases proceed through the formation of a phosphoenzyme. Four different types of residues have been shown to be implicated in covalently binding phosphate in these enzymes: serine, as in alkaline phosphatase [25]; histidine, as in acid phosphatase [26], fructose 2,6-bisphosphatase [27] and phosphoglycerate mutase [28]; cysteine in tyrosine phosphatases [29]; and aspartate in ATPases of the Na^+/K^+ ATPases family [30,31]. Histidine and cysteine are most likely not involved since there is no conserved cysteine or histidine in L-3-phosphoserine phosphatase (see Fig. 1). Furthermore the lability of the phosphoenzyme to alkali and to NH_2OH indicates an acyl-phosphate linkage, as in phosphoaspartate or phosphoglutamate. The low stoichiometry of phosphorylation and the lability of the phosphoenzyme bond prevented us from isolating a tryptic phosphopeptide suitable for sequencing. The identity of the phosphorylated residue remains therefore to be determined.

Two highly conserved motifs containing aspartate residues are found in L-3-phosphoserine phosphatases: DXDST and GDGXXD. The first one shares two residues (DXXXT) with the consensus phosphorylation site of ATPases of the Na^+/K^+ ATPase family. The second motif is also found in the same family of ATPases, where it is highly conserved (e.g., residues 713 to 718 of the alpha subunit of human Na^+/K^+ ATPase). These considerations indicate that there may be functional homology between phosphoserine phosphatases and ATPases.

Acknowledgements: The authors thank H.G. Hers for critical reading of the manuscript, B.J. Van den Eynde for providing the human cDNAs, G. Berghenouse for technical assistance and H. Degand (Laboratory of Physiological Biochemistry, University of Louvain) for performing the microsequencing. This work was supported by the Actions de Recherche Concertées and by the Belgian Federal Service for Scientific, Technical and Cultural affairs. M.H.R. is Chercheur Qualifié and M.V.D.C., Chargé de Recherches of the Belgian Fonds National de la Recherche Scientifique.

References

[1] L.F. Borkenhagen, E.P. Kennedy, Biochim. Biophys. Acta 28 (1958) 222–223.
[2] F.C. Neuhaus, W.L. Byrne, Biochim. Biophys. Acta 28 (1958) 223–224.
[3] J. Jaeken, M. Detheux, L. Van Maldergem, J.-P. Frijns, P. Alliet, M. Foulon, H. Carchon, E. Van Schaftingen, J. Inherit. Metab. Dis. 19 (1996) 223–226.
[4] A.F. Neuwald, G.V. Stauffer, Nucl. Acid Res. 13 (1985) 7025–7039.
[5] L.F. Borkenhagen, E.P. Kennedy, J. Biol. Chem. 234 (1959) 849–853.
[6] L.M. Cagen, H.C. Friedmann, J. Biol. Chem. 247 (1972) 3382–3392.
[7] G. Lennon, C. Auffray, M. Polymeropoulos, M.B. Soares, Genomics 33 (1996) 151–152.
[8] B. Gaugler, N. Brouwenstijn, V. Vantomme, J.P. Szikora, C.W. Van der Spek, J.J. Patard, T. Boon, P. Schrier, B.J. Van den Eynde, Immunogenetics 44 (1996) 323–330.
[9] M. Hérin, C. Lemoine, P. Weynants, F. Vessière, A. Van Pel, A. Knuth, R. Devos, T. Boon, Int. J. Cancer 39 (1987) 390–396.
[10] F.W. Studier, B.A. Moffatt, J. Mol. Biol. 189 (1986) 113–130.
[11] M. Veiga-da-Cunha, M. Detheux, N. Watelet, E. Van Schaftingen, Eur. J. Biochem. 225 (1994) 43–51.
[12] J. Jaeken, M. Detheux, L. Van Maldergem, M. Foulon, H. Carchon, E. Van Schaftingen, Arch. Dis. Child. 74 (1996) 542–545.
[13] D.A. Hager, R.R. Burgess, Anal. Biochem. 109 (1980) 76–86.
[14] M.H. Rider, M. Puype, J. Van Damme, K. Gevaert, S. De Boeck, J. D'Alayer, H.H. Rasmussen, J.E. Celis, J. Vandekerckhove, Eur. J. Biochem. 230 (1995) 258–265.
[15] C.H. Fiske, Y.P. Subbarow, J. Biol. Chem. 66 (1925) 375–400.
[16] R.W. McClard, Anal. Biochem. 96 (1979) 500–503.
[17] M.M. Bradford, Anal. Biochem. 72 (1976) 248–254.
[18] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. 215 (1990) 403–410.
[19] W.F. Bridgers, J. Biol. Chem. 242 (1967) 2080–2085.
[20] J.D. Corbin, E.M. Reimann, Methods Enzymol. 38 (1974) 287–290.
[21] J.J. Witt, R. Roskoski Jr., Anal. Biochem. 66 (1975) 253–258.
[22] U.K. Laemmli, Nature 227 (1970) 680–685.
[23] B.R. Nechay, L.B. Nanninga, P.S.E. Nechay, R.L. Post, J.J. Grantham, I.G. Macara, L.F. Kubena, T.D. Phillips, F.H. Nielsen, Fed. Proc. 45 (1986) 123–132.
[24] V. Shetty, K.T. Shetty, Neurochem. Res. 15 (1990) 1203–1210.
[25] L. Engström, Biochim. Biophys. Acta 52 (1961) 49–59.
[26] M. Igarashi, H. Takahashi, N. Tsuyama, Biochim. Biophys. Acta 220 (1970) 85–92.
[27] S.J. Pilkis, M. Walderhaug, K. Murray, A. Beth, S.D. Venkataramu, J. Pilkis, M.R. El-Maghrabi, J. Biol. Chem. 258 (1983) 6135–6141.
[28] Z.B. Rose, Arch. Biochem. Biophys. 140 (1970) 508–513.
[29] K.L. Guan, J.E. Dixon, J. Biol. Chem. 266 (1991) 17026–17030.
[30] L.E. Hokin, P.S. Sastry, P.R. Galsworthy, A. Yoda, Proc. Natl Acad. Sci. USA 54 (1965) 177–184.
[31] R.L. Post, S. Kume, J. Biol. Chem. 248 (1973) 6993–7000.

The Plant Journal (1998) **16**(4), 443–452

Molecular characterization of plastidic phosphoserine aminotransferase in serine biosynthesis from *Arabidopsis*

Chai-Ling Ho, Masaaki Noji, Maiko Saito, Mami Yamazaki and Kazuki Saito*
Laboratory of Molecular Biology and Biotechnology in Research Center of Medicinal Resources, Faculty of Pharmaceutical Sciences, Chiba University, Inage-ku, Yayoi-cho 1–33, Chiba 263–8522, Japan

Summary

Serine biosynthesis in plants proceeds by two pathways; a photorespiratory pathway which is associated with photorespiration and a pathway from phosphoglycerate. A cDNA encoding plastidic phosphoserine aminotransferase (PSAT) which catalyzes the formation of phosphoserine from phosphohydroxypyruvate has been isolated from *Arabidopsis thaliana*. Genomic DNA blot analysis indicated that this enzyme is most probably encoded by a single gene and is mapped on the lower arm of chromosome 4. The deduced protein contains an N-terminal extension exhibiting the general features of a plastidic transit peptide, which was confirmed by subcellular organelle localization using GFP (green flourescence protein). Northern analysis indicated preferential expression of PSAT in roots of light-grown plants, supporting the idea that the phosphorylated pathway may play an important role in supplying the serine requirement of plants in non-green tissues. *In situ* hybridization analysis of PSAT revealed that the gene is generally expressed in all types of cells with a significantly higher amount in the meristem tissue of root tips.

Introduction

Serine is a key intermediate in a number of important metabolic pathways including the photorespiratory metabolism of glycolate to phosphoglycerate (Walton and Woolhouse, 1986). It is also important as a precursor biomolecule in the generation of glycine, tryptophane, cysteine and in the interconversion of homocysteine and methionine (Ireland and Hiltz, 1995; Walton and Woolhouse, 1986). Besides these functions, serine is involved in the synthesis of phospholipids, porphyrins, purines and thymidine, and is the source of one-carbon units.

In the yeast *Saccharomyces cerevisiae*, serine and glycine are mainly synthesized by two pathways. On fermentable carbon sources, serine is generated via the glycolytic pathway from phosphoglycerate, whereas on non-fermentable carbon sources, the other pathway, namely the gluconeogenic pathway (Melcher *et al.*, 1995) starting from glyoxylate, is utilized. In animals and bacteria, serine is mainly synthesized via the phosphorylated pathway (Ichihara and Greenberg, 1957) utilizing phosphoglycerate, which is derived from glycolysis or from the oxidative or reductive pentose phosphate pathway.

The situation is more complex in photosynthetic plant tissues since biosynthesis of serine can proceed via the photorespiratory pathway (Tolbert, 1980) (Figure 1). Glycolate is the substrate utilized during photorespiration (Hatch, 1976). Its production in the chloroplasts initiates the photorespiratory pathway where glycolate is oxidized to glyoxylate from which glycine is generated. In the mitochondria, serine is generated by reactions involving two molecules of glycine; one molecule is oxidatively decarboxylated and deaminated to produce an active hydroxymethyl group that becomes attached to the β-carbon of the second molecule. The glycine decarboxylase multi-enzyme complex (GDC), along with the enzyme serine hydroxymethyltransferase (SHMT), is responsible for the respiratory conversion of glycine to serine (Neuberger *et al.*, 1986; Srinivasan and Oliver, 1995). The cDNAs encoding for the four different component enzymes of GDC (Kim and Oliver, 1990; Macherel *et al.*, 1990; Turner *et al.*, 1992b, 1992c) and SHMT (Turner *et al.*, 1992a, 1993) from plants have been cloned and characterized. Most of the serine produced is returned to Calvin cycle via hydroxypyruvate and glycerate.

In the absence of photorespiration, serine could be synthesized from phosphoglycerate by the non-phosphorylated pathway, involving glycerate and hydroxypyruvate as intermediates or via the series of reactions catalyzed by phosphoglycerate dehydrogenase, phosphoserine aminotransferase (PSAT) and phosphoserine phosphatase (Bryan, 1988). In animal cells, the non-phosphorylated pathway is postulated to participate in serine catabolism rather than serine biosynthesis (Snell, 1986). However, the conversion of serine to phosphoglycerate is readily reversible in plants (Tolbert, 1980). Serine is probably formed by this pathway in germinated pea cotyledons which lacked phosphoserine aminotransferase activities (Walton and Woolhouse, 1986).

In the phosphorylated pathway, the glycolytic intermediate, phosphoglycerate, is oxidized by phosphoglycerate dehydrogenase to form phosphohydroxypyruvate,

Received 3 August 1998; revised 28 September 1998; accepted 1 October 1998.
*For correspondence (fax +81 43 2902905; e-mail ksaito@p.chiba-u.ac.jp).

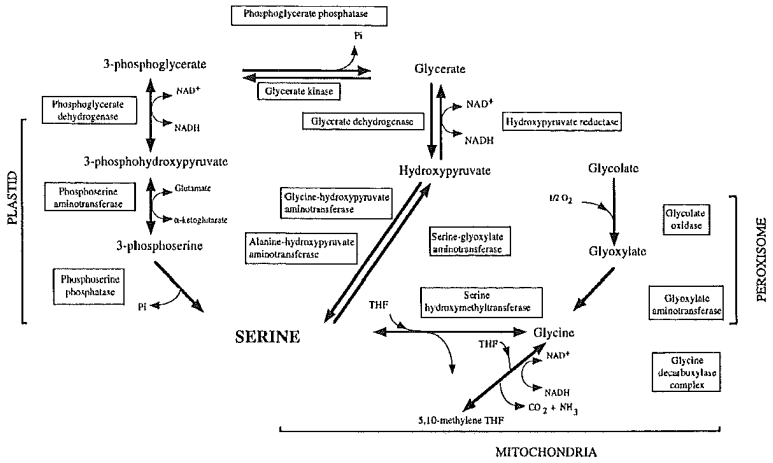


Figure 1. Main routes involved in the biosynthesis of serine in higher plants. THF, tetrahydrofolate.

which in turn is transaminased by phosphoserine aminotransferase to yield phosphoserine. In the final step, dephosphorylation of phosphoserine is catalyzed by phosphoserine phosphatase (Ichihara and Greenberg, 1957; Stolz and Dörnemann, 1994). It has been proposed that the latter pathway is probably of minor significance compared to the photorespiratory pathway during daylight hours where photorespiration takes place, but may be of more importance in the dark or early in leaf development (Ireland and Hiltz, 1995). However, there was evidence for a phosphorylated pathway operating in green tissues during photosynthesis (Ireland and Hiltz, 1995; Keys, 1980).

The existence of metabolites that can be formed by two or more pathways raises interesting questions concerning their regulation and expression. Since various biosynthetic enzymes are distributed among different subcellular organelles in plant cells, it is possible that significant regulation of the metabolism of serine may control the rate and extent of the transport of intermediates and products of the pathways between various metabolic sites (Bryan, 1988).

In our previous paper (Saito *et al.*, 1997), we reported the isolation of a cDNA clone encoding PSAT from spinach and the preliminary characterization of the cDNA. In the present paper, we describe for the first time cDNA and genomic cloning, biochemical characterization and expression as well as the immunolocalization and *in situ* hybridization of PSAT in *Arabidopsis thaliana*.

Results

Isolation and sequence of cDNA encoding PSAT of Arabidopsis

The Arabidopsis ESTs clone FAFL52, which shows 55–67% homology to bacterial and yeast PSAT proteins, was used to screen 2.5×10^5 plaques from a whole plant *Arabidopsis* cDNA library constructed in λgt11. Among the four positive clones selected for further studies, CPSAT-5 was shown to contain the largest cDNA insert (1.7 kb).

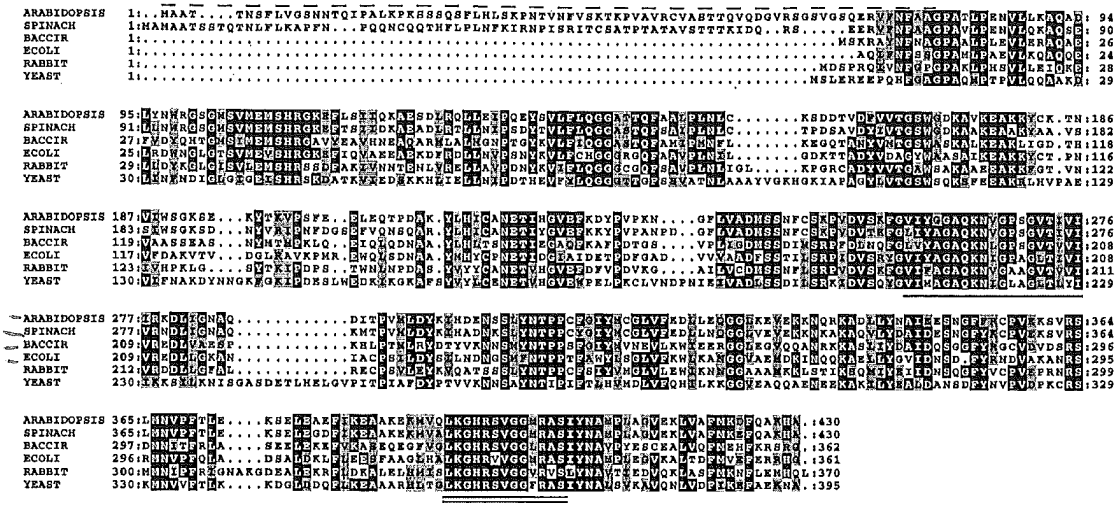
Sequence analysis revealed the presence of an open reading frame which is 1290 nucleotides in length. The deduced 430 amino acids encode PSAT and have a molecular weight of 47 359 Da. The first ATG triplet is designated as the translation start site because the surrounding sequence (AATCATGGC) agrees well with the favoured sequence (AACAATGGC) flanking the consensus functional plant-initiator codon (Lutcke *et al.*, 1987). Furthermore, an in-frame stop codon is found 51 nucleotides upstream of the translation start methionine in the cDNA sequence. Primer extension experiments confirmed the above assumption (see below). The cDNA sequence shows the presence of a 3′-untranslated region of 189 nucleotides downstream of the translation stop codon. The polyadenylation signal is most likely the AATAAT motif located 105 nucleotides downstream from the translation stop codon.

The deduced amino acid sequence shares 34–71% homology with bacterial, yeast and spinach PSATs (Figure 2a). Both the signature sequence of PSAT proteins (Belhumeur *et al.*, 1994) and the conserved sequence motif involving the pyridoxal-phosphate binding lysine site found in Class V aminotransferase (Ouzounis and Sander, 1993) are conserved in *Arabidopsis* PSAT. A phylogenetic tree (Figure 2b) indicates that PSATs from *Arabidopsis* and spinach form a family distinct from bacterial and animal PSATs.

Genomic cloning and sequence analysis of PSAT

On screening approximately 2×10^5 plaques of the *Arabidopsis* genomic library with CPSAT-5, 2 positive clones, GPSAT-28 and GPSAT-45, were selected for further analysis. The 5 kb fragment, which covers part of the PSAT structural gene and the upstream region, was isolated from

(a)



(b)

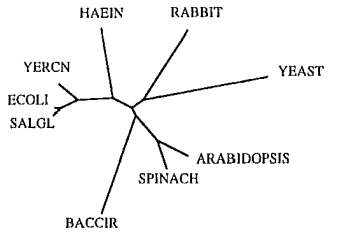


Figure 2. Comparison of deduced amino acid sequences of PSAT from *Arabidopsis* and other organisms.

YEAST, PSAT from *Saccharomyces cerevisiae* (Belhumeur *et al.*, 1994); HAEIN, *Haemophilus influenza* (Fleischmann *et al.*, 1995); SALGL, *Salmonella gallinarum* (Griffin, 1990); ECOLI, *Escherichia coli* (Duncan and Coggins, 1986); YERCN, *Yersinia enterocolitica* (O'Gaora *et al.*, 1989); RABBIT, *Oryctolagus cuniculus* (Misrahi *et al.*, 1987); SPINACH, *Spinacia oleracea* (Saito *et al.*, 1997); BACCIR, *Bacillus circulans* (Battchikova *et al.*, 1996); ARABIDOPSIS, *Arabidopsis thaliana* (the present study).

(a) Multiple alignment of deduced amino acid sequences of PSAT from *Arabidopsis* and other organisms. Underlined letters indicate the conserved sequence motif involving the pyridoxal-phosphate binding lysine site found in Class-V aminotransferase (Ouzounis and Sander, 1993). Double-underlined letters indicate the signature sequence for PSAT proteins (Belhumeur *et al.*, 1994). A dashed line indicates the putative transit peptide used for construction of GFP fusion protein.

(b) Phylogenetic tree of PSATs using the phylogenetic program package. Phylogenetic tree, constructed using PHYLIP 3.57c program by J. Felsenstein, indicates that PSATs from *Arabidopsis* and spinach form a distinct family from bacterial, yeast and animal PSATs.

clone GPSAT-45 and sequenced, whilst the 1.1 kb fragment which covers the rest of the structural gene and the 3′ region was isolated from GPSAT-28. Comparison of this sequence with the cDNA sequence of CPSAT-5 revealed that no intron is present in the coding region of PSAT gene. The transcriptional start site was mapped by using primer extension. A single transcriptional start site is located at 171 bp before the translation start site. An AT-rich sequence is located at –12 to –22 and a potential CAAT sequence at –94 to –97.

Southern blot analysis was used to estimate the number of genes encoding PSAT in *Arabidopsis*. *Arabidopsis* genomic DNA digested with six restriction endonucleases was fractionated on an agarose gel. The result, as shown in Figure 3, revealed that a single major fragment hybridized upon digestion with *Eco*RV, *Sac*I and *Xba*I. However, upon digestion by *Bam*HI, *Bgl*II and *Eco*RI, several bands were observed. The appearance of the strongly hybridizing bands is due to the presence of several restriction sites for the endonucleases in the genomic sequence, whereas the weak signals observed are most probably due to the other isoforms of PSAT, or are maybe just due to the presence of pseudogenes. These results suggested that *Arabidopsis* contains a single gene corresponding to the isolated cDNA and a few related sequences homologous to the cDNA in the genome. Using the 10.5 kb fragment from GPSAT-45 as probe for analysis, the *Arabidopsis* PSAT gene was mapped to the lower part of chromosome 4.

Biochemical characterization of recombinant PSAT

The 60 amino acid leader sequence exhibits the general features of a transit peptide for transportation of protein to the plastid. It starts with methionine-alanine; is rich in the hydroxylated amino acids, serine and threonine (16/60); has the small hydrophobic amino acids, alanine and valine (12/60); is essentially deficient in the acidic amino acids, aspartic acid and glutamic acid (1/60); and has a net positive charge (pI = 10.7). The prediction from the PSORT program (GenomeNet service, Osaka University) was that it is localized in the stroma of the chloroplast.

The fusion protein of CPSAT-GFP (green fluorescence protein) could be detected in intact tissues after delivering the constructs into *Arabidopsis* leaves by particle gun bombardment. The expressed signals in both constructs containing the predicted transit peptides of PSAT from *Arabidopsis* (Figure 4b) and spinach (Saito *et al.*, 1997) (data not shown) were similar to those exhibited by transit peptide of the ribulose-1,5-bisphosphate carboxylase small subunit polypeptide of *Arabidopsis* (Krebbers *et al.*, 1988) (Figure 4a) which was already known to be sufficient for translocation into chloroplasts (Chiu *et al.*, 1996). The bright green fluorescence was clearly visible despite the red chlorophyll autofluorecence from chloroplasts under blue-light.

These results confirmed that the N-terminal sequence of *Arabidopsis* and spinach PSATs are sufficient for translocalization of passenger protein into chloroplasts. Thus, it is likely that the PSATs are plastidic proteins.

The identity of the isolated cDNA, CPSAT-5, was confirmed by successful complementation of a serine-auxotrophic mutant lacking the PSAT locus *Ser*C. The mutant *E. coli* KL282 (*tonA22, phoA4*(Am) (*serS14, serC16*), *serS13, ompF627, supD32* (*serUI32*), *relA1, pit10, spoT1*, T_2R) (Low *et al.*, 1971) was transformed with an expression plasmid, pPSAT-AB13TS, in which the expression of PSAT cDNA is regulated by the *lac*Z promoter. Transformants could grow in the M9 minimal medium in the absence of serine, whereas the pTV118N transformed *E. coli* KL282 was not able to grow without supplementation with serine (Figure 5), confirming the authenticity of the CPSAT-5 encoding the functional PSAT.

The recombinant PSAT was overexpressed in *E. coli* AD494 using a pET32a(+) vector system with a strong T7 promoter. The His-tagged PSAT protein was purified by using a His-Bind^R purification column (Novagen). Enzyme activity was not detected in the purified protein which was in the insoluble and inactive form. Expression of PSAT in a soluble form was too low to be detected by using SDS-PAGE, however it exhibited PSAT activity of 0.72 ± 0.15 μmoles min^{-1} mg of crude $protein^{-1}$. Specific activity of PSAT overexpressed in *E. coli* BL21, using the pET3d vector system, without His-tag was also investigated. The protein in the crude extract exhibited higher activity (1.54 ± 0.01 μmoles min^{-1} mg of crude $protein^{-1}$) than the overproduced protein in the *E. coli* AD494/

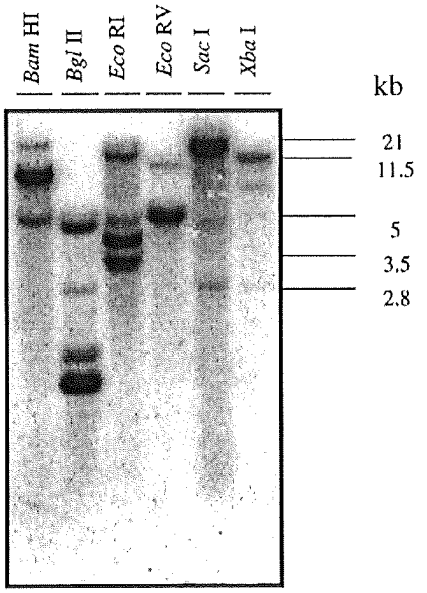


Figure 3. Southern blot analysis of genomic DNA.
Genomic DNA was extracted from the leaves of 3-week-old seedlings. For every lane, about 5 μg genomic DNA was digested with restriction enzyme, separated by electrophoresis through a 0.8% (w/v) agarose gel, transferred to a Hybond N^+ membrane (Amersham), and then hybridized with ^{32}P-labelled CPSAT-5. The final washing was performed in 0.1 × SSPE, 0.1% SDS at 65°C for 10 min.

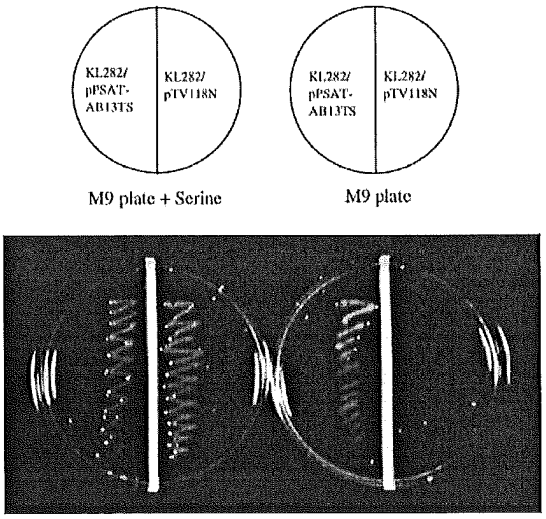


Figure 5. Functional complementation of *serC*$^-$ *E. coli* KL282 (Low *et al.*, 1971) by transformation with the expression vector pPSAT-AB13TS carrying CPSAT-5.
pTV118N is the empty vector used as a negative control. The transformed bacteria were cultured on M9 minimal agar plates with 1.0 mg ml^{-1} serine (left plate) or without serine (right plate) supplementation at 30°C.

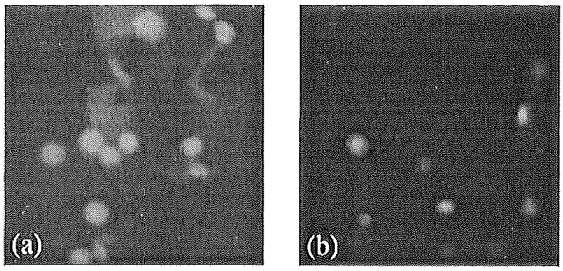


Figure 4. Organelle sublocalization of PSAT using GFP as reporter protein. Particle gun bombardment was carried out using Helios Gene-Gun System, at pressure 100 psi to transform 3-week-old *Arabidopsis* seedlings. A 228-bp (dashed line in Figure 2a) PCR-amplified fragment of PSAT from *Arabidopsis* was subcloned into plasmid CaMV35S-sGFP(S65T)-NOS3'. The 35SΩ-TP-sGFP(S65T) construct, carrying the transit peptide sequence obtained from the ribulose-1,5-bisphosphate carboxylase small subunit polypeptide of *Arabidopsis* (Krebbers *et al.*, 1988) was used as a positive control. (a) Fluorescence signal pattern exhibited by 35SΩ-TP-sGFP(65T), which was known to be targeted to chloroplasts (Chiu *et al.*, 1996).
(b) Fluorescence signal pattern exhibited by CPSAT-TP of *Arabidopsis* (this study). Bar = 100 μm.

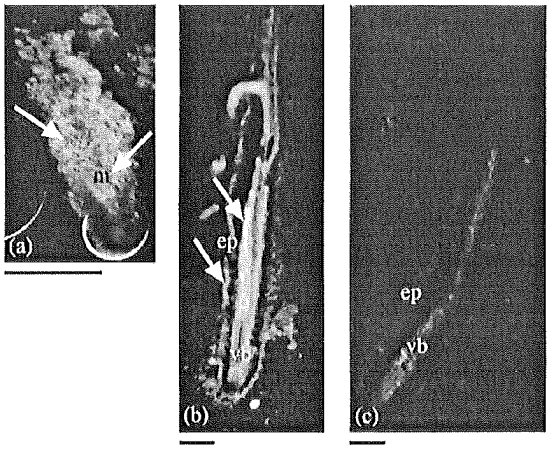


Figure 8. *In situ* hybridization of *Arabidopsis* with PSAT antisense probe. Blue color is toluidine blue staining. Longitudinal sections of root tip (a) and root (b,c) under dark field optics. (a,b) Signals = bright spots indicated by arrowheads. (c) Negative control, being hybridized with PSAT sense probe. Bars = 200 μm; m, meristem; ep, epidermis; vb, vascular bundle.

pET32a(+) system. The recombinant protein was detected as a 41 kDa protein by SDS-PAGE (Figure 6). The PSAT expressed was truncated to remove approximately 60 amino acid residues of putative transit peptide from the full-length deduced open reading frame. Yet it was catalytically active and, in fact, resulted in enhanced protein accumulation in *E. coli* crude extract compared to the full-length deduced open reading frame.

We determined K_m values for glutamate and phosphohydroxypyruvate using the overexpressed crude proteins of PSAT from *Arabidopsis* and spinach in *E. coli* BL21/pET3d system. Double reciprocal plots of the initial rates data demonstrated K_m values for the *Arabidopsis* PSAT of 70 μM and 5 mM for glutamate and phosphohydroxypyruvate, respectively, and for the spinach PSAT of 150 μM and 3 mM for glutamate and phosphohydroxypyruvate, respectively. In the range of 5–50 mM, serine, threonine, valine, glycine, tryptophane and O-acetyl-L-serine had no effect on the rate of reaction. Enzyme inhibition was observed with a high concentration of cysteine.

The PSAT activity of recombinant enzyme of *Arabidopsis* was inhibited by the antiserum raised against the His-Bind^R purified recombinant protein overexpressed in the *E. coli* AD494/pET32a(+) system. Inhibition was linear with the amount of antibodies added to the reaction mixture (data not shown).

Gene expression of PSAT

Northern blot analysis was carried out to examine the level of mRNA expression in shoot and root tissues from light-grown and dark-treated plants (Figure 7). The highest level of PSAT mRNA expression was observed in the light-grown roots followed by a significant amount of mRNA expression in light-grown shoots. A lower level of expression was also detected in dark-grown shoots and roots. The preferential expression of PSAT mRNA in the roots was contrasted with mRNAs of two enzymes involved in the photorespiratory pathway, the H-subunit of GDC (Srinivasan and Oliver, 1995) and serine hydroxymethyltransferase (SHMT) (Turner *et al.*, 1992a), which accumulated primarily in the light-grown shoot tissues, although a less pronounced signal was also detected in the dark-grown shoots.

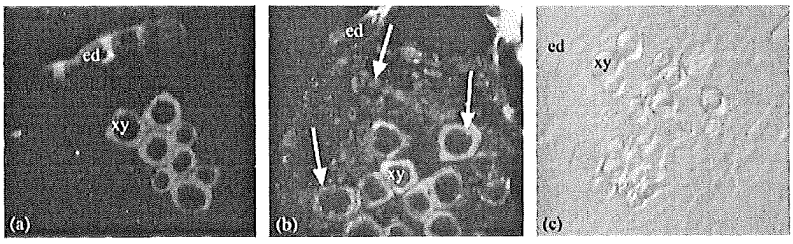


Figure 10. Immunolocalization of PSAT protein in *Arabidopsis* (a–c are cross-sections of stem).
(a) Negative control, incubated with pre-immune serum, in which only autofluorescence signals of xylem were observed.
(b) Incubated with rabbit antiserum raised against CPSAT, green fluorescence signals (indicated by arrowheads) were detected in the stele, preferentially associated with xylem.
(c) A non-fluorescent micrograph of stem. ed, endodermis; xy, xylem. Bar = 100 μm.

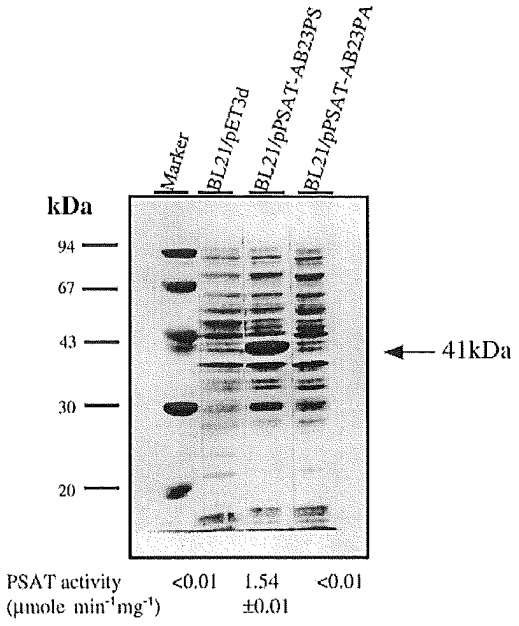


Figure 6. Overexpression of CPSAT-5 in pET3d and determination of enzyme activity of PSAT in the crude extract of *E. coli* BL21.
Lanes: M, standard proteins as molecular weight marker (Pharmacia); BL21/pET3d, soluble fraction of *E. coli* BL21 carrying an empty vector pET3d; BL21/pPSAT-AB23PS, soluble fraction of *E. coli* BL21 carrying plasmid pPSAT-AB23PS with CPSAT-5 in sense orientation to the promoter; BL21/pPSAT-AB23PA, soluble fraction of *E. coli* carrying plasmid pPSAT-AB23PA with CPSAT-5 in antisense orientation. For the enzyme assay, the activities of PSAT were determined in soluble protein extracts using the method described previously (Duncan and Coggins, 1986). Data are expressed as mean ± SD (n = 3) μmoles min^{-1} mg of crude $protein^{-1}$.

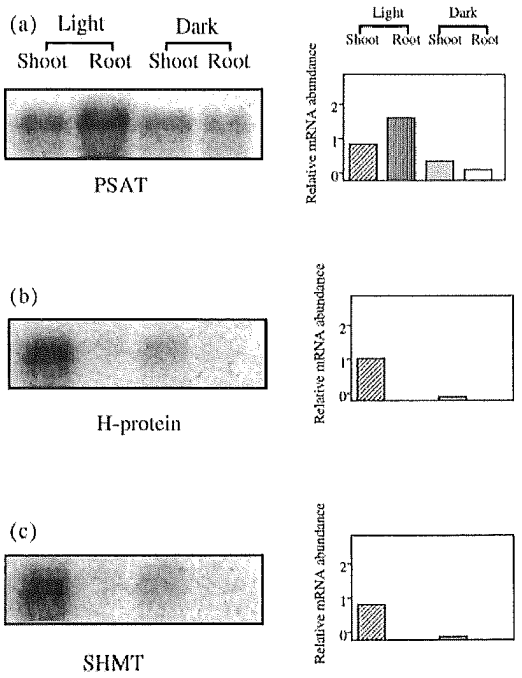


Figure 7. Northern analysis of RNA from different tissues of *Arabidopsis* seedlings.
Ten μg total RNA was separated under denaturing conditions in a 1.2% agarose gel containing formaldehyde and transferred to a Hybond N^+ membrane (Amersham) and then probed with ^{32}P-labelled cDNA clone. The final washing was performed in 0.1 × SSPE, 0.1% SDS at 65°C for 10 min.
(a) The mRNA of PSAT was accumulated in both leaf and root tissues, with the highest amount preferentially expressed in root tissues from light-grown plants.
(b,c) Preferential expression of H-protein, a subunit of GDC, and SHMT mRNA in leaf tissues of light-grown plants although a much lower amount of expression was also detected in dark-treated leaf tissues.

In situ hybridization results suggested that mRNA of PSAT was expressed at some level in all types of cells in leaf, stem and root tissues (data not shown). However, a strong signal was detected in meristem tissue of root tips (Figure 8a), indicating a significantly higher level of expression in fast proliferating tissues. The root tissues surrounding the vascular bundle also seemed to be sites of preferential expression (Figure 8b).

Protein accumulation and immunolocalization of PSAT

The antibodies raised against the overexpressed PSAT in the pET32a(+) system recognized the overexpressed protein in pET3d as well as the associated protein in *Arabidopsis*. Western blot results (Figure 9) showed that PSAT protein was undetectable in dark-treated seedlings. Immunolocalization experiments (Figure 10) indicated that PSAT protein was preferentially accumulated in the stele, in particular the cells close to the xylem in leaf, stem and root sections.

Discussion

Structural and functional properties of the PSAT molecule

Our previous paper (Saito *et al.*, 1997) described the preliminary characterization of PSAT by cDNA cloning from spinach. The present study is the first extensive molecular characterization of the serine biosynthetic pathway from phosphoglycerate in plants.

The nucleotide sequence corresponding to the *Arabidopsis* PSAT mRNA encodes a precursor of mature protein of 430 amino acid residues. The spinach and *Arabidopsis* PSAT enzymes were engineered so that the protein expressed in *E. coli* had approximately 60 amino acid residues truncated from the full-length deduced open reading frame. These enzymes were catalytically active. Therefore, the N-terminal extension of these cDNAs may function as a transit peptide. Essential common characters of a chloroplast transit peptide suggests that the PSAT is located in the plastid/chloroplast. The gene product shows

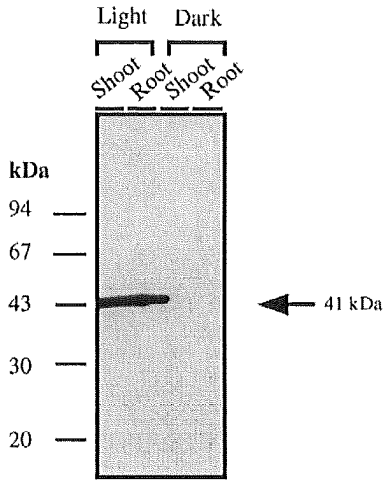


Figure 9. Western blot analysis of PSAT.
Protein extracted from 3-week-old *Arabidopsis* was separated by SDS-PAGE and electrochemically transferred onto Immobilon P membrane. It was then reacted with 1:500 dilution of rabbit antiserum raised against purified recombinant PSAT fused with His-tag. The immuno-reactive proteins were visualized using phosphatase-labelled goat antirabbit and NBT and BCIP as substrates. PSAT protein was preferentially expressed in both leaves and roots of *Arabidopsis* light-grown seedlings. No significant signal was detected for the extracts from dark-treated seedlings.

high homology to PSAT amino acid sequences from other sources. Both the signature sequence for PSAT proteins and the common sequence motif involving the pyridoxal binding lysine found in Class V aminotransferase are conserved in *Arabidopsis* PSAT. The gene product functionally complemented an *E. coli serC* mutant strain. Enzyme activity associated with PSAT in *E. coli* crude extract, which overexpressed the cDNA, further confirmed its identity.

Serine failed to inhibit PSAT from both *Arabidopsis* and spinach. Our findings agree with observations by Reynolds *et al.* (1988) using pea PSAT but were in contrast to the report of Larsson and Albertsson (1979) that PSAT from spinach was strongly inhibited by serine. The antibodies against the recombinant PSAT could inhibit the activity of the overexpressed PSAT *E. coli* crude extract.

Physiological significance of gene expression and accumulation of PSAT protein

The co-existence of two or more biosynthetic pathways in higher plant suggests that either of two pathways may be preferentially utilized depending upon the physiological state of plants at a given time. It is conceivable that serine is derived from glycolate via glycine during photosynthesis. In the absence of photorespiration, in particular, in rapidly proliferating tissues of low photosynthetic activity (Cheung *et al.*, 1968; Reynolds and Blevins, 1986), the phosphorylated pathway may play the primary role in the total cellular supply of serine. This is probably because the photorespiratory pathway does not efficiently operate in the dark (Keys, 1980). Therefore, photosynthetic tissues, e.g. leaves of light-grown plants, may have a greater capacity for utilizing glycine as a serine precursor; whereas in non-photosynthetic tissues, e.g. roots, dark-treated leaves and etiolated leaves, etc., the phosphorylated pathway may take over to be the main supply of serine in plants (Ireland and Hiltz, 1995; Reynolds *et al.*, 1988; Saito *et al.*, 1997).

Under normal conditions, only minor PSAT activities were measured in the leaves of spinach (Larsson and Albertsson, 1979) and pea (Walton and Woolhouse, 1986), while in tissues associated with rapid cell proliferation, e.g. seed leaves and apical meristems of pea (Cheung *et al.*, 1968) as well as root tissues of soybean and lupin (Reynolds and Blevins, 1986), considerable amounts of PSAT were found. Stolz and Dörnemann (1994) also detected PSAT activity in a *Scenedesmus obliquus* mutant with low photorespiratory activity under normal CO_2 condition which is comparable to tissues with rapid cell proliferation. The phosphorylated pathway enzymes are important in proliferating plant and animal tissues where there is a high serine requirement. Snell (1985) reported that the rat neoplastic tissues have higher serine hydroxymethyl transferase and phosphoserine aminotransferase activities, perhaps reflecting the increased demand for one-carbon fragments ultimately required for DNA synthesis. Our findings are consistent with the previous reports. We provide molecular evidence that mRNA of PSAT is preferentially expressed in roots from light-grown plants (Figure 7), especially in the meristem tissue as revealed by *in situ* hybridization results (Figure 8). A minor amount of PSAT mRNA expression in the shoots from light-grown plants may be due to the non-photosynthetic cell types that co-exist in the leaves and stems.

In contrast to the preferential expression of PSAT, mRNA abundance of H-protein (a subunit of GDC) and SHMT, the two enzymes which are responsible for the photorespiratory pathway in shoots of light-grown plants, far exceeded their expression in roots or dark-grown shoots (Figure 7). Again, this finding supports the idea that the photorespiratory pathway predominates in photosynthetic tissues. The phosphorylated pathway may play an important role in supplying serine to the root. However, further physiological studies of the fluxes through the two pathways have to be compared to confirm the above result. Further investigations are necessary to clarify the interaction of the two pathways at a molecular and physiological level precisely.

Immunolocalization of PSAT protein (Figure 10) provided further information that PSAT protein is mainly detected in cells in and around the vascular bundle, and is preferentially associated with xylem in leaf, stem and root tissues. Since these cells do not take part in photosynthesis, the phosphorylated pathway may therefore play

an important role in supplying serine requirement to these tissues. It appears likely that the increase in protein synthesis is partly regulated at the mRNA level.

Meanings of plastidic localization of PSAT

Reynolds and Blevins (1986) reported that PSAT was localized in the proplastid, a conclusion based on enzyme isolation from the plastidic fraction. However, Larsson and Albertsson (1979) detected 90% of PSAT activity outside the chloroplasts in leaves in contrast to phosphoserine phosphatase, the enzyme responsible for the final step in the phosphorylated pathway which was of exclusively chloroplastic origin. Thus, Larsson and Albertsson (1979) proposed that the phosphohydroxypyruvate formed by phosphoglycerate dehydrogenase was exported from chloroplasts to form phosphoserine which was then imported into the chloroplasts to serve as a substrate for phosphoserine phosphatase. We provide the first subcellular localization evidence which shows the targeting of spinach and *Arabidopsis* PSAT proteins into leaf chloroplasts. The function of the transit-peptides located at the N-terminus of the cDNAs were confirmed.

The Southern blot analysis suggested the presence of a single gene corresponding to the isolated cDNA, but we do not exclude the possibility of related sequences encoding for other isoforms of PSAT in *Arabidopsis* since several weakly hybridizing bands were also observed in addition to the expected bands based on the genomic DNA sequence. However, it is premature to suggest the isolated gene is a main gene without first knowing whether other genes encoding PSAT species exist, and how many, as well as their physiological roles. The presence of another isoform of PSAT which is not localized in the plastid remains a possibility and requires further investigation.

Experimental procedures

Plant materials

Arabidopsis thaliana ecotype Columbia seeds were germinated and grown on GM agar plates (Valvekens *et al.*, 1988) under 16 h/8 h light and dark cycles at 22°C for 3 weeks. For the dark-treated seedlings used for Northern analyses, 2-week-old seedlings were wrapped in aluminium foil and subsequently grown for another 1 week before RNA extraction was carried out.

Isolation of cDNA and genomic clones

For the isolation of cDNA which encodes for PSAT, approximately 2.5×10^5 amplified plaques of the λgt11 cDNA library of *A. thaliana* ecotype Columbia were subjected to screening with the ^{32}P-labelled probe synthesized from the cDNA insert of the *Arabidopsis* EST, FAFL52 (accession no. ATT4557). The *Arabidopsis* cDNA clone, CPSAT-5, was used to screen for genomic clones from an *Arabidopsis* EMBL3 SP6/T7 library (CLONTECH). Approximately 2×10^5 amplified plaques were screened.

Hybridization of the membranes (Hybond N^+ Amersham) was carried out at 65°C in 5 × SSPE (0.9 M NaCl, 0.05 M sodium phosphate pH 7.7, 5 mM EDTA), 0.5% SDS, 5 × Denhardt's solution and 25 μg ml^{-1} salmon sperm DNA. Membranes were washed at maximum stringency in 1 × SSPE, 0.1% SDS for 10 min twice and final washing was conducted at 65°C in 0.1 × SSPE, 0.1% SDS for 10 min, and then exposed to Fuji X-ray film.

Nucleic acid preparation and blot analyses

Genomic DNA was extracted from the leaves of 3-week-old seedlings as described by Dellaporta *et al.* (1983). Isolation of total RNA was performed by a modified guanidine-HCl method as described by Sambrook *et al.* (1989) from the leaves and roots of 3-week-old seedlings. DNA and RNA blots were hybridized with ^{32}P-labelled probe synthesized from cDNA insert fragment of the isolated clone. To investigate the mRNA expression levels of GDC and SHMT, ^{32}P-labeled probes synthesized from cDNA inserts of *Arabidopsis* EST clones, 200K16T7 and 111M16T7, respectively, were used. To verify equivalent loadings of RNA on blots, membranes were hybridized with a ^{32}P-labelled probe synthesized from rice rDNA (pRR217). Relative values of mRNA were calculated based on the hybridization intensities of specific signals on blots quantified by a BAS-2000 image analyzer (Fuji).

Restriction fragment length polymorphism (RFLP) mapping was carried out by 30 recombinant inbred lines (Lister and Dean, 1993). Hybridization and washing were carried out as described above except that the final washing condition was reduced to 0.5 × SSPE, 0.1% SDS for 10 min at 65°C. ^{32}P-labelled probe was synthesized from genomic clone, GPSAT-45, and used for hybridization. The map distance was calculated based on the RFLP profiles generated by *Hha*I.

Overexpression and purification of recombinant enzyme

A general method of DNA engineering was followed according to Sambrook *et al.* (1989). *Nco*I sites were created on both ends of the coding region by polymerase chain reaction (PCR) engineering using synthetic primers: AB1 (5′-GTCTACCATGGGCTCCGTCGGATCCCAAG-3′) and AB2 (5′-CACGTCCATGGAGAAAGATTTTGAT-3′). The engineered DNA fragment was inserted into the *Nco*I site of pET3d or pET32a(+) (Novagen-TaKaRa, Kyoto, Japan), in which the cDNA was placed under a strong Ø10 promoter in both sense and antisense orientation. The plasmids were then introduced into *E. coli* BL21 (DE3) pLysS or AD494 (DE3) pLysS in which the gene for lysogenic T7 RNA polymerase under *lac*UV5 promoter is induced by isopropyl-1-thio-β-D-galactopyranoside (IPTG). The recombinant protein accumulated in the insoluble fraction of *E. coli* AD494 extracts was purified using HIS-BIND^R buffer kit (Novagen) following the instructions as described in the manual.

Transformants of *E. coli* KL282 (*tonA22, phoA4*(Am) (*serS14, serC16*), *serS13, ompF627, supD32* (*serUl32*), *relA1, pit10, spoT1, T_2R*) mutant (Low *et al.*, 1971) with plasmids pTV118N and pCPSAT-AB13T, respectively, were tested for their ability to grow on M9 minimal medium.

Protein harvesting from Arabidopsis

Approximately 0.1 g of leaf tissue was added to 300 μl of lysis buffer (2 mM Tris–HCl, pH 8.0; 0.14 M NaCl; 1 mM phenylmethyl-

sulfonyl fluoride) and homogenized with a pestle in an Eppendorf tube, followed by centrifugation at 10 000 **g** for 5 min at 4°C. The supernatant was used for protein quantitation and SDS-PAGE.

Enzyme assay

The assay for PSAT activity is based on that described by Duncan and Coggins (1986). Assay mixtures contained 50 mM Tris–HCl (pH 8.2), 32 mM ammonium acetate, 2 mM glutamate, 0.2 mM NADH, 2.5 mM phosphohydroxypyruvate, 2 units of glutamate dehydrogenase and 10 µl protein extract. The reaction at 25°C was initiated by the addition of phosphohydroxypyruvate, and the absorbance change at 340 nm was measured.

Subcellular organelle localization of PSAT by GFP

To verify the subcellular localization of PSAT in plant cells, a 228 bp PCR-amplified fragment of PSAT from *Arabidopsis* using primers 5′-AGTGTGTTCGACATGGCGGCTACGACGAACT-3′ and 5′-CCTGGCCATGGCAAGTTGAAGACACGTTCT-3′; and a 216 bp PCR-amplified fragment of PSAT from spinach using primers 5′-GAGAAGTCGACATGGCAATGGCGGCCACCT-3′ and 5′-CCTGGCCATGGCAAAGTTGAAGACGCGTT-3′, were subcloned into plasmid CaMV35S-sGFP(S65T)-NOS3′ (Chiu *et al.*, 1996) with a CaMV 35S-promoter at its 5′-end, GFP as reporter gene and 3′NOS as transcription terminator. Plasmid construct, 35SΩ-TP-sGFP(S65T) (Chiu *et al.*, 1996) carrying a transit peptide sequence obtained from the ribulose-1,5-bisphosphate carboxylase small subunit polypeptide of *Arabidopsis* (Krebbers *et al.*, 1988), was used as a positive control. These plasmids were used for subsequent particle gun bombardment.

Particle gun bombardment was carried out using the Helios Gene-Gun System (Bio-Rad) following standard protocol provided by the supplier. Plants were incubated for 20 h under illumination conditions at 22°C after bombardment. Individual leaves were viewed with a fluorescent microscope (BX50-FLA, Olympus) using the Chroma's Dual Band filter, FITC & TRITC (Olympus). Photographs were taken using Fujichrome ISO 100 film.

Immunolocalization of PSAT

Immunolocalization of PSAT protein was performed on 3-week-old *A. thaliana* ecotype Columbia seedlings. The plant tissues were initially fixed in 4% formaldehyde, 50% ethanol and 5% acetic acid overnight at 4°C. Fixed tissues were dehydrated and embedded in polyethyleneglycol (PEG)#1540 (Kanto Chemical) as described by Marrison and Leech (1992). Sections of 8 µm were mounted onto slides coated with 3-aminopropyltriethoxysilane (Sigma) and left to dry overnight at 42°C. The sections were soaked in phosphate-buffered saline (PBS; 0.16 M NaCl, 8 mM Na_2HPO_4, 2.7 mM KCl, 1.5 mM KH_2PO_4) to remove PEG and incubated overnight at 4°C with 100 µl of 1 : 200 dilution of rabbit antiserum raised against PSAT in 0.5% (w/v) BSA/PBS. After sequential washing with 0.5% (w/v) BSA/PBS for 15 min, with 0.01% (v/v) Tween 20/PBS for 15 min and with PBS for 15 min, sections were incubated with FITC conjugated goat antirabbit IgG (Sigma) in a diluted concentration recommended by the supplier, in 0.5% (w/v) BSA/PBS at room temperature for 2 h. Slides were mounted with a drop of Vectashield and viewed by a fluorescent microscope (BX50-FLA, Olympus) with Chroma's Dual Band filter, FITC & TRITC. Photographs were taken using Fujichrome ISO 400 film.

In situ *hybridization*

In situ hybridization was carried out on 3-week-old *A. thaliana* seedlings, which were fixed in 4% formaldehyde, 50% ethanol and 5% acetic acid for 3 h at room temperature. Fixed tissues were dehydrated and embedded in paraffin wax according to standard procedures (Angerer and Angerer, 1992). Ten µm sections were mounted onto slides coated with 3-aminopropyltriethoxysilane and pre-treated for hybridization according to Angerer and Angerer (1992). ^{35}S-UTP-labelled sense and antisense RNA probes were generated by run-off transcription with T7 and T3 RNA polymerase (Promega). Labelled RNA probes were hydrolyzed to an average length of 300 nucleotides. The hybridization mix contained ^{35}S-labelled mRNA (5× 10^6 cpm slides^{-1}), 10 mM Tris–HCl (pH 7.5), 50% (v/v) formamide, 0.3 M NaCl, 1 mM EDTA, 150 µg ml^{-1} yeast tRNA, 1 × Denhardt's, 10% (w/v) dextran sulfate, and 70 mM dithiothreitol. RNase treatment washing steps were performed as described by Angerer and Angerer (1992). Slides were coated with Hypercoat emulsions (Amersham) and then developed with Kodak D-19. After staining with 0.05% (w/v) toluidine blue and following dehydration, the slides were mounted in PolyMount (Polysciences Inc., Warrington, USA). Photographs were taken in a microscope (BX50, Olympus) using dark-field optics and Fujichrome ISO 100 film.

Miscellaneous techniques

DNA subcloning and sequencing, sodium dodecyl sulfate (SDS)-polyarylamide gel electrophoresis (PAGE), Western blot analysis, protein quantitation and primer extension were carried out as described by Sambrook *et al.* (1989).

Acknowledgements

We thank ABRC of Ohio State University for *Arabidopsis* ESTs clones 200K16T7 and 111M16T7 and Dr Y. Parmentier (Institut de Biologie Moleculaire des Plantes) for *Arabidopsis* ESTs clone FAFL52. We would also like to thank Dr M. Arnold (Nottingham Arabidopsis Stock Centre) for kindly mapping the genes; Dr B.J. Bachmann (*E. coli* Genetic Stock Center, Department of Biology, Yale University) for the gift of *E. coli* KL282; and Dr Y. Niwa (Shizuoka Prefecture University, Shizuoka, Japan) for the gift of plasmids CaMV35S-sGFP(S65T)-NOS3′ and 35S-TP-sGFP(S65T). We are indebted to Mr H. Hoshi for assistance with the gene sequencing and Ms A. Watanabe for preparing the gene mapping filter membranes. This research was supported in part by grants-in-aids for scientific research from Ministry of Education, Science, Sports and Culture, Japan.

References

Angerer, L.M. and Angerer, R.C. (1992) *In situ* hybridization to cellular RNA with radiolabelled RNA probes. In *In Situ Hybridization: A Practical Approach, The Practical Approach Series.* (Wilkinson, D.G., ed.). Oxford: IRL Press, pp. 15–32.

Battchikova, N., Himanen, J.P., Ahjolahti, M. and Korpela, T. (1996) Phosphoserine aminotransferase from *Bacillus circulans* subsp. alkalophilus: purification, gene cloning and sequencing. *Biochim. Biophys. Acta*, **1295**, 187–194.

Belhumeur, P., Fortin, N. and Clark, M.W. (1994) A gene from *Saccharomyces cerevisiae* which codes for a protein with significant homology to the bacterial 3-phosphoserine aminotransferase. *Yeast*, **10**, 385–389.

Bryan, J.K. (1988) Advances in the biochemistry of amino acid

biosynthesis. In *The Biochemistry of Plants. A Comprehensive Treatise*, Vol. 16 (Miflin, B.J. and Lea, P.J. eds.). New York: New York Academic Press, pp. 175–176.

Cheung, G.P., Rosenblum, I.Y. and Sallach, H.J. (1968) Comparative studies of enzymes related to serine metabolism in higher plants. *Plant Physiol.* **43**, 1813–1820.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. and Sheen, J. (1996) Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325–330.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) *A Plant DNA Minipreparation, Version II. Plant Mol. Biol. Rep.* **1**, 19–23.

Duncan, K. and Coggins, J.R. (1986) The *ser*C-*aro*A operon of *Escherichia coli*: a mixed function operon encoding enzymes from two different amino acid biosynthetic pathways. *Biochem. J.* **234**, 49–57.

Fleischmann, R.D., Adams, M.D., White, O. *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenza* Rd. *Science*, **269**, 496–512.

Griffin, H.G. (1990) Nucleotide sequence of the *Salmonella ser*C gene. *Nucl. Acids Res.* **18**, 4260.

Hatch, M.D. (1976) Photosynthesis: the path of carbon. In *Plant Biochemistry*, 3rd edn (Borner, J. and Varner, J.E., eds). New York: Academic Press Inc, pp. 797–844.

Ichihara, A. and Greenberg, D.M. (1957) Further studies on the pathway of serine formation from carbohydrates. *J. Biol. Chem.* **224**, 331–340.

Ireland, R.J. and Hiltz, D.A. (1995) Glycine and serine synthesis in non-photosynthetic tissues. In *Amino Acids and Their Derivatives in Higher Plants* (Wallsgrove, R.M., ed.). Cambridge: Cambridge University Press, pp. 111–118.

Keys, A.J. (1980) Synthesis and interconversion of glycine and serine. In *The Biochemistry of Plants*, Vol. 5 (Stumpf, P.K. and Conn, E.E., eds.). New York: Academic Press Inc, pp. 359–374.

Kim, Y. and Oliver, D.J. (1990) Molecular cloning, transcriptional characterization, and sequencing of cDNA encoding the H-protein of the mitochondrial glycine decarboxylase complex in peas. *J. Biol. Chem.* **265**, 848–853.

Krebbers, E., Seurinck, J., Herdies, L., Cashmore, A.R. and Timko, M.P. (1988) Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis*. *Plant Mol. Biol.* **11**, 745–759.

Larsson, C. and Albertsson, E. (1979) Enzymes related to serine synthesis in spinach chloroplasts. *Physiol. Plant.* **45**, 7–10.

Lister, C. and Dean, C. (1993). Recombinant imbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.* **4**, 745–750.

Low, B., Gates, F., Goldstein, T. and Soll, D. (1971) Isolation and partial characterization of temperature-sensitive *Escherichia coli* mutants with altered leucyl- and seryl-transfer ribonucleic acid synthetases. *J. Bacterial.* **108**, 742–750.

Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6**, 43–48.

Macherel, D., Lebrum, M., Gagnon, J., Neuberger, M. and Douce, R. (1990) cDNA cloning, primary structure and gene expression for H-protein, a component of the glycine-cleavage system (glycine decarboxylase) of pea (*Pisum sativa*) leaf mitochondria. *Biochem. J.* **268**, 783–789.

Marrison, J.L. and Leech, R.M. (1992) Co-immunolocalization of topoisomerase II and chloroplast DNA in developing, dividing and mature wheat chloroplasts. *Plant J.* **2**, 783–790.

Melcher, K., Rose, M., Kunzlen, M., Braus, G.H. and Entian, K.D. (1995) Molecular analysis of the yeast SER1 gene encoding 3-phosphoserine aminotransferase: regulation by general control and serine repression. *Curr. Genet.* **27**, 501–508.

Misrahi, M., Atger, M. and Milgrom, E. (1987) A novel progesterone-induced messenger RNA in rabbit and human endometria. Cloning and sequence analysis of the complementary DNA. *Biochemistry*, **26**, 3975–3982.

Neuberger, M., Bourguignon, J. and Douce, R. (1986) Isolation of a large complex from the matrix of the pea leaf mitochondria involved in the rapid transformation of glycine into serine. *FEBS Lett.* **20**, 718–722.

O'Gaora, P., Maskell, D., Coleman, D., Cafferkey, M. and Dougan, G. (1989) Cloning and characterization of the *ser*C and *aro*A genes of *Yersinia enterolitica*, and construction of an *aro*A mutant. *Gene*, **84**, 23–30.

Ouzounis, C. and Sander, C. (1993) Homology of the NifS family of proteins to a new class of pyridoxal phosphate-dependent enzymes. *FEBS Lett.* **322**, 159–164.

Reynolds, P.H. and Blevins, D.G. (1986) Phosphoserine aminotransferase in soybean root nodules. *Plant. Physiol.* **81**, 283–293.

Reynolds, P.H., Hine, A. and Rodber, K. (1988) Serine metabolism in legume nodules: purification and properties of phosphoserine aminotransferase. *Physiol. Plant.* **74**, 194–199.

Saito, K., Takagi, Y., Ho, C.-L., Takahashi, H. and Noji, M. (1997) Molecular cloning, characterization and expression of cDNA encoding phosphoserine aminotransferase involved in phosphorylated pathway of serine biosynthesis from spinach. *Plant Mol. Biol.* **33**, 359–366.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Snell, K. (1985) Enzymes of serine metabolism in normal and neoplastic rat tissues. *Biochim. Biophys. Acta*, **843**, 276–281.

Snell, K. (1986) The duality of pathways for serine biosynthesis is a fallacy. *Trends Biochem. Sci.* **11**, 241–243.

Srinivasan, R. and Oliver, D.J. (1995) Light-dependent and tissue-specific expression of the H-protein of the glycine decarboxylase complex. *Plant Physiol.* **109**, 161–168.

Stolz, M. and Dörnemann, D. (1994) Purification, characterization and N-terminal sequence of phosphoserine aminotransferase from the green alga *Scenedesmus obliquus*, mutant C-2A[1]. *Z. Naturforsch.* **49c**, 63–69.

Tolbert, N.E. (1980) Photorespiration. In *The Biochemistry of Plants*, Vol. 2. (Davies, D.D., ed.). New York: Academic Press, pp. 488–525.

Turner, S.R., Hellens, R., Ireland, R., Ellis, N. and Rawsthorne, S. (1993) The organisation and expression of the genes encoding the mitochondrial glycine decarboxylase complex and serine hydroxymethyltransferase in pea (*Pisum sativuum*). *Mol. Gen. Genet.* **236**, 402–408.

Turner, S.R., Ireland, R., Morgan, C. and Rawsthorne, S. (1992a) Identification and localization of multiple forms of serine hydroxymethyltransferase in pea (*Pisum sativum*) and characterization of a cDNA encoding a mitochondrial isoform. *J. Biol. Chem.* **267**, 13528–13534.

Turner, S.R., Ireland, R. and Rawsthorne, S. (1992b) Cloning and characterization of the P subunit of glycine decarboxylase from pea (*Pisum sativum*). *J. Biol. Chem.* **267**, 5355–5360.

Turner, S.R., Ireland, R. and Rawsthorne, S. (1992c) Purification and primary amino acid sequence of the L subunit of glycine decarboxylase. *J. Biol. Chem.* **267**, 7745–7750.

Valvekens, R.D., Van Montagu, M. and Van Lijsebettens, M. (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl Acad. Sci. USA*, **85**, 5536–5540.

Walton, N.J. and Woolhouse, H.W. (1986) Enzymes of serine and glycine metabolism in leaves and non-photosynthetic tissues of *Pisum sativum* L. *Planta*, **167**, 119–128.

GenBank database accession numbers D88541 (PSAT cDNA) and AB010408 (PSAT genomic).

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 11, Issue of March 13, pp. 6066–6073, 1998
Printed in U.S.A.

Serine Hydroxymethyltransferase Is Maternally Essential in *Caenorhabditis elegans**

(Received for publication, November 3, 1997, and in revised form, December 23, 1997)

Greg P. Vatcher‡, Colin M. Thacker§, Titus Kaletta¶‖, Heinke Schnabel¶‖, Ralf Schnabel¶‖, and David L. Baillie‡

From the ‡Institute of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, the §Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada, the ¶Max-Planck-Institut für Biochemie, Martinsried D-82152, Federal Republic of Germany, and the ‖Institut für Genetik, Braunschweig D-38023, Federal Republic of Germany

The *mel-32* gene in the free living soil nematode *Caenorhabditis elegans* encodes a serine hydroxymethyltransferase (SHMT) isoform. Seventeen ethylmethanesulfonate (EMS)-induced mutant alleles of *mel-32*(SHMT) have been generated, each of which causes a recessive maternal effect lethal phenotype. Animals homozygous for the SHMT mutations have no observable mutant phenotype, but their offspring display an embryonic lethal phenotype. The Mel-32 phenotype has been rescued with a transgenic array containing only *mel-32*(SHMT) genomic DNA. Heteroduplex analysis of the 17 alleles allowed 14 of the mutations to be positioned to small regions. Subsequent sequence analysis has shown that 16 of the alleles alter highly conserved amino acids, while one allele introduces a stop codon that truncates two thirds of the predicted protein. *mel-32*(SHMT) has a 55–60% identity at the amino acid level with both isoforms of SHMT found in yeast and humans and a 50% identity with the *Escherichia coli* isoform. The *C. elegans mel-32* mutation represents the first case where SHMT has been shown to be an essential gene.

Serine hydroxymethyltransferase (SHMT)[1] is a highly conserved, ubiquitous, pyridoxal 5′-phosphate (PLP)-containing enzyme that catalyzes the reversible conversion of serine and tetrahydrofolate to glycine and 5,10-methylenetetrahydrofolate (1). The β-carbon of serine is the major source of one-carbon units in the one-carbon metabolic pool (2). The methyl group passed on through the folate cofactor of SHMT is used in thymidylate, methionine, lipid, and purine biosynthesis. SHMT also catalyzes many secondary reactions, such as amino acid transaminations (for a full review of the reactions catalyzed by SHMT, see Ref. 1).

Eukaryotic organisms have two nuclear genes that encode two distinct isoforms of this enzyme, a cytosolic form and a mitochondrial form. Recent evidence suggests that there is a subcellular partitioning of the reactions catalyzed by SHMT, with the serine to glycine conversion occurring in the mitochondria, and the glycine to serine conversion occurring in the cytoplasm (3). A comparison of the determined and predicted amino acid sequences of many SHMTs reveals a striking degree of homology between all the known forms of this enzyme (2, 4, 5). These alignments show that the amino and carboxyl termini of SHMT are less conserved but that the middle two thirds of the sequence have long stretches of very high identity.

Serine hydroxymethyltransferase levels are elevated in rapidly proliferating cell lines and tumors (6, 7). When lymphocytes are treated with a mitogenic stimulus, the enzymatic activity of SHMT is increased and the incorporation of the β-carbon of serine into DNA is increased (6). SHMT activity is also increased during the S phase of the cell cycle, suggesting that a product of SHMT activity is utilized during cell division. Conversely, when cells stop proliferating, SHMT levels decrease. Retinoic acid treatment of P19 embryonal carcinoma cells stops proliferation and stimulates differentiation. Furthermore, it has been shown recently that retinoic acid causes a 50% decrease in SHMT transcript levels (7). Based on these findings, serine hydroxymethyltransferase has been proposed as a potential chemotherapy target (8). A chemical or drug that decreases SHMT activity may cause rapidly proliferating tumor cells to quiesce.

Prior to this study, the only observable phenotype of SHMT deficiency was glycine auxotrophy. In *Escherichia coli*, when the single copy of SHMT (*glyA*) is mutated, glycine auxotrophy results (9). In the yeast *Saccharomyces cerevisiae*, glycine auxotrophy is observed only when both forms of SHMT and a third gene, *glyA*, are mutated (4) suggesting a functional redundancy in the glycine synthesis pathway. In addition, a line of Chinese hamster ovary cells that lack mitochondrial SHMT activity are glycine auxotrophs (3, 10). This deficiency is rescued when the cells are transfected with a cDNA encoding the human mitochondrial SHMT (11).

This report presents the cloning and characterization of a *Caenorhabditis elegans* SHMT homolog called *mel-32*. Mutations in *mel-32* result in a maternal effect lethal (Mel) phenotype. The *C. elegans mel-32* mutation is the first reported case of a serine hydroxymethyltransferase deficiency causing lethality.

EXPERIMENTAL PROCEDURES

Growth and Handling of C. elegans—All strains of *C. elegans* were grown at 20 °C on nematode growth medium plates streaked with *E. coli* OP50 as a food source. Standard genetic manipulation followed previously described protocols (12).

Computer Analysis—The analysis of sequence data, sequence comparisons, and data base searches were performed with ACeDB (a *C. elegans* data base) (13),[2] the BLAST[3] (provided by the NCBI server) and FASTA programs (14, 15), CLUSTALW (16), and MacDNASIS Pro

* This work was supported by a Natural Sciences and Engineering Research Council of Canada grant (to D. L. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** To whom correspondence should be addressed. Tel./Fax: 604-291-4597; E-mail: dbaillie@darwin.mbb.sfu.ca.

[1] The abbreviations used are: SHMT, serine hydroxymethyltransferase; PLP, pyridoxal 5′-phosphate; EMS, ethylmethanesulfonate; *mel*, maternal effect lethal; bp, base pair(s); PCR, polymerase chain reaction.

[2] ACeDB documentation, code, and data are available from anonymous FTP servers at lirmm.lirmm.fr, cele.mrc-lmb.cam.ac.uk, and ncbi.nlm.nih.gov.

[3] The BLAST program is available via the World Wide Web (http://www.nchi.nlm.nih.gov/blast).

(Hitachi Software Engineering Co., Ltd.). PCR primers were designed with the aid of Oligo (17).

Cloning the C. elegans SHMT Gene—The cosmid clone of C05D11 (GenBank accession number U00048) was kindly provided by Dr. Alan Coulson (Medical Research Council, Hinxton, United Kingdom). A computer restriction analysis of C05D11 (MacDNASIS) revealed that coding elements 11 (SHMT) and 13 could be isolated as a 5052-base pair (bp) *Pvu*II-*Sca*I (Life Technologies, Inc., Pharmacia Biotech Inc.) fragment (residues 36311–41363 of the GenBank entry). The plasmid pGV9 (Fig. 2.) was constructed by ligating the *Pvu*II-*Sca*I fragment into *Eco*RV (Life Technologies, Inc.)-cut pBluescript II KS$^+$ (Stratagene). The plasmid pC05.11 (Fig. 2.) was constructed by cutting pGV9 with *Pst*I (New England Biolabs) and religating, leaving a 3196-bp fragment of C05D11 (residues 36311–39507), containing only open reading frame 11 (SHMT), in pBluescript.

Mutant Rescue—The subclones containing the SHMT gene were injected into the syncytial gonad of adult wild-type (N2) *C. elegans* hermaphrodites together with the dominant marker *rol-6(su1006)* (18, 19) contained on plasmid pCes1943.[4] Stable transgenic strains, those expressing the roller phenotype in successive generations, were used in the rescue experiments.

Heteroduplex Analysis—DNA from individual *C. elegans* heterozygous for each of the SHMT mutations was isolated as described by Barstead *et al.* (20) with the modifications of Williams *et al.* (21). PCR using *Taq* polymerase (BioCan Scientific) was performed as follows; 5 pmol of each of five sets of primers (synthesized by DNAgency, Malvern, PA) were used (Table I). The samples were incubated at 94 °C for 1 min on an Idaho Technology 1605 Air Thermo-Cycler before commencement of 30 cycles of amplification (94 °C for 10 s, 59 °C (62 °C for the D primer set) for 20 s, and 72 °C for 40 s). Following a 2-min incubation at 72 °C, 1.25 μl of EDTA (0.1 M, pH 7.5) was added to terminate the reaction. Heteroduplex analysis was performed using a mutation detection enhancement gel matrix (J.T. Baker) following the method of Nijbroek *et al.* (22) with the following changes; a heteroduplex denaturation/reannealing profile of 95 °C for 3 min, 85 °C for 3 min, 75 °C for 5 min, 65 °C for 5 min, 55 °C for 5 min, and 37 °C for 5 min was used. The heteroduplex DNA was then resolved on a 0.5× mutation detection enhancement gel at 400 V for 16–20 h.

DNA Sequence Analysis—DNA from individual nematodes homozygous for the *mel-32*(SHMT) mutations was isolated and amplified with the same primers used in the heteroduplex analysis. A total of 0.5 μl of this template DNA was incubated at 95 °C for 3 min on a Precision Scientific GTC-2 Genetic Thermo Cycler before commencement of 35 cycles of amplification (94 °C for 45 s, 59 °C (62 °C for the D primer set) for 30 s, and 72 °C for 1 min) followed by a polishing step of 72 °C for 7 min. The PCR products from two separate reactions were pooled, purified by agarose gel electrophoresis, and collected with a Qiagen QIAquick gel extraction kit. A total of 100–200 ng of each PCR product was sequenced on both strands using FS *Taq* terminator chemistry (Applied Biosciences) on a Perkin-Elmer GeneAmp PCR System. The reactions were run on an Applied Biosciences model 373A automated DNA sequence analyzer located at the Nucleic Acid-Protein Service Unit, University of British Columbia.

RESULTS AND DISCUSSION

Identification of a C. elegans Homolog of SHMT—A search of the ACeDB program revealed one SHMT homolog, C05D11.11, located on chromosome III within the area defined by the cosmid C05D11. The position of the CeSHMT gene was based on both a Genefinder[5] prediction and on five partial cDNAs isolated as part of the *C. elegans* genome sequencing project. The predicted gene contains four exons coded in 1599 bp of genomic sequence. The predicted 484-amino acid protein sequence was used to search the GenBank data bank with the FASTA algorithm (14, 15), and the closest homologs were rabbit and human cytosolic SHMTs (61.7% and 61.3% identity, respectively, in a 470-amino acid overlap).

A sequence comparison of SHMTs shows the highly conserved nature of this protein and reveals many conserved domains (Fig. 1). The amino- and carboxyl-terminal 50 amino acids are the least conserved, and the central three-quarters of the protein contain large stretches of completely conserved amino acids. The PLP cofactor binding site (residues 301–305 in the *C. elegans* protein) and the active site lysine (residue 306) are conserved in every case.

Rescue of mel-32 with C. elegans SHMT—Our laboratory is in the process of constructing a transgenic library of sequenced cosmids that can be used for high resolution genetic mapping (23). Stable transgenic arrays are generated, which can act as cosmid sized duplications in rescue experiments. If the genomic DNA present in the extrachromosomal array rescues the recessive lethality of an essential gene then the wild-type copy of the mutation must be present in the DNA defined by the transgenic. Previous rescue experiments in our laboratory (23) have placed five essential genes in the genomic region defined by the cosmid C05D11. The genes *let-713, let-721, let-725, let-756*, and *mel-32* are all rescued by C05D11, making these genes candidates for potential mutations in CeSHMT.

The *mel-32* mutant phenotype was rescued with CeSHMT genomic DNA, indicating that *mel-32* encodes SHMT. All rescue experiments were performed with the canonical allele of *mel-32, s2518*, which was isolated in an EMS mutagenesis screen for maternal effect lethals.[6] Subsequently, *mel-32* was found to be allelic with 16 EMS-induced Mel alleles from a collection of maternal effect embryonic lethal mutations on chromosome III isolated by H. Schnabel and R. Schnabel.[7] Hermaphrodites homozygous for the *mel-32* mutations have no observable mutant phenotype, but their self-fertilized offspring display an embryonic lethal phenotype and arrest at about the 100 cell stage. The Mel phenotype of *mel-32* was rescued with pGV9, which contains coding element 11 (SHMT) and 13 (Fig. 2). To determine in which of these two genes the mutations reside, a smaller subclone, pC05.11 (Fig. 2), containing only gene 11 (SHMT) was constructed. This subclone, pC05.11, gave a partial rescue of the Mel-32 phenotype. From *mel-32*;pC05.11 transgenic hermaphrodites, a small number of progeny hatch and grow to adulthood but are not themselves fertile. pGV9 contains the entire 1254 bp of intergenic sequence between gene 11 (CeSHMT) and gene 13, while pC05.11 contains only 416 bp upstream of the CeSHMT ATG start codon. The partial rescue by pC05.11 suggests that some important regulatory sequences are missing from the smaller subclone.

Heteroduplex Analysis of mel-32(SHMT) Mutations—The 17 alleles of *mel-32* were analyzed by heteroduplex analysis to define more precisely the regions where the mutations occurred. Five sets of overlapping PCR products were generated from animals heterozygous for each allele. The primers used for PCR are listed in Table I, the exact location of each is shown in Fig. 3, and the overlap of each product is shown in Fig. 4. Each amplified DNA product was heat-denatured and allowed to cool slowly. Any PCR product that contains a mutation will have a mixture of homoduplexes and heteroduplexes, paired wild-type and mutant strands that contain base pair mismatches. These duplexes were run on a special mutation detection enhancement gel (see "Experimental Procedures"), which can reveal a single base pair mismatch in a short strand of DNA (24). The presence of multiple bands indicates a mutation in that particular overlapping fragment. Using this procedure, 13 out of the 17 mutations were placed into specific regions of the gene.

These mutations were distributed as follows. Region A (Fig. 4) contained one mutation: *t1473* (which was also detected in region B; see below). Region B contained six mutations: *t1473* (see below), *t1555, t1597, t1665, t1679*, and *s2518*. Region C contained three mutations: *t1552, t1616*, and *t1631*. Region D

[4] S. J. M. Jones and W. B. Barbazuk, personal communication.
[5] P. Green and L. Hillier, manuscript in preparation.
[6] G. Vatcher, unpublished results.
[7] H. Schnabel and R. Schnabel, unpublished results.

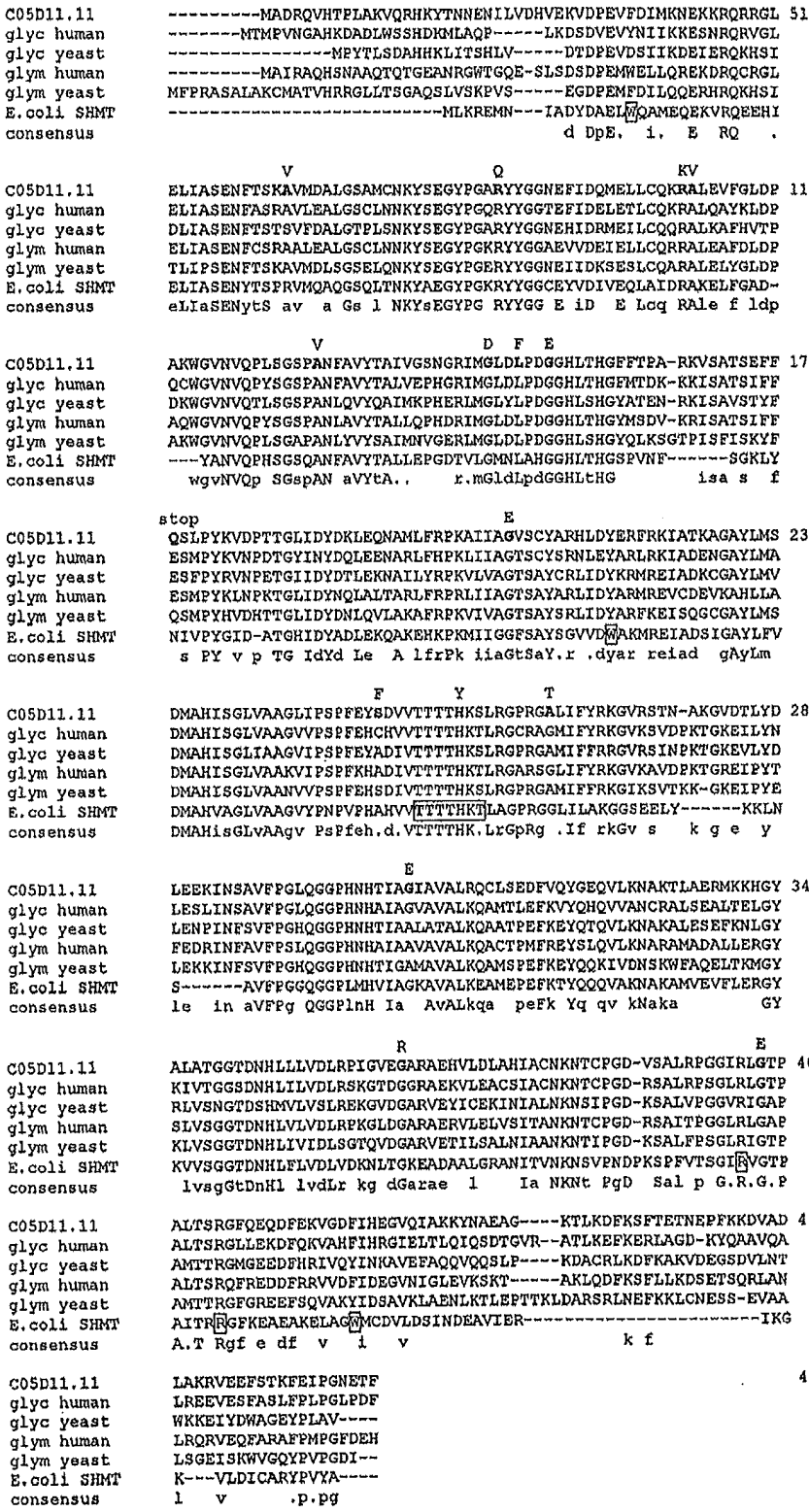


FIG. 1. **Comparison of the amino acid sequences of selected SHMTs.** The aligned sequences are *C. elegans mel-32* (*C05D11.11*), human cytosolic (*glyc human*) and mitochondrial (*glym human*), *S. cerevisiae* cytosolic (*glyc yeast*) and mitochondrial (*glym yeast*), and *E. coli GlyA* (*E.coli*

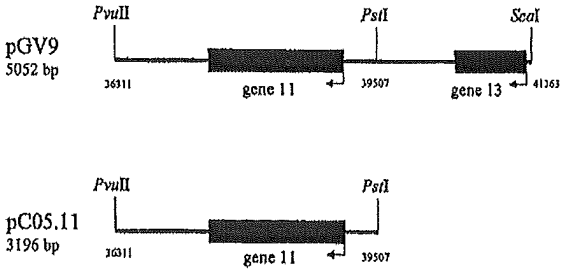


FIG. 2. **SHMT subclones.** The genomic subclones used to rescue *mel-32* are shown. Nucleotide numbers correspond to the GenBank entry for C05D11 (U00048). Restriction sites used for subcloning are indicated.

TABLE I
PCR primers

Primer	Sequence	Product length
		bp
A1	ATGCATATACGTGTTGCC	399
A2	TAATACCGAGCTCCAGGG	
B1	GAAGCGGCAGCGTCGTGG	408
B2	GAGACTTTGCGAGCTGGG	
C1	TATGGGCCTCGATCTTCC	483
C2	GTGTCAACTCCCTTTGCG	
D1	CAACGACGCACAAGTCGC	449
D2	GAGACGAATGCCTCCTGG	
E1	CGGAGTTGAAGGAGCTCG	474
E2	AAGTGATAAAATGGCGGG	

contained three mutations: *t1456*, *t1474*, and *t1576*. Region E contained a single mutation: *t1632*. For allele *t1666*, no mutations were detected in regions A, C, D, and E, but a single PCR product could not be obtained for region B. It was therefore assumed that the lesion in this allele was contained in region B, which was later confirmed by direct sequence analysis. The position of the molecular lesions in alleles *t1520*, *t1607*, and *t1671* could not be detected using this procedure.

Sequence Analysis of mel-32(SHMT)—The same oligonucleotide primers used to generate PCR products for the heteroduplex analysis (Table I) were used to amplify DNA from homozygous mutant animals from the region in which the mutation for each allele was detected. For alleles *t1520*, *t1607*, and *t1671*, in which heteroduplexes were not observed, DNA from all five regions was amplified from mutant strains. The PCR products were then directly sequenced in both orientations (see "Experimental Procedures").

The sequence analysis results are summarized in Table II and Fig. 4. Allele *t1631* contains the only nonsense mutation identified for *mel-32*. The codon for glutamine 171 is mutated into an amber stop codon. Any truncated protein produced in this mutant would lack the terminal two thirds of the enzyme, including the active site. This allele is assumed to be a null. The phenotype of this mutant is indistinguishable from the other 16 alleles, suggesting that all of the alleles may be null, or that all reduce the SHMT activity below the threshold required for embryonic survival.

Alleles *t1456* and *t1474* both contain the same base pair mutation, which changes glycine 313 to glutamic acid. In allele *t1616*, alanine 268 is changed to threonine. The amino acids at positions 313 and 268 of the SHMT consensus sequence (Fig. 1) are always either a glycine or an alanine. It is assumed that replacing a small, neutral amino acid with a larger acidic or polar one warps the secondary structure of SHMT enough to disrupt its enzymatic function.

Eukaryotic forms of the SHMT holoenzyme exist as tetramers of identical subunits, and the predicted secondary structure contains alternating helices and strands, placing SHMT in the α/β class of proteins (5). Fig. 1 reveals that there are 22 completely conserved glycines in the SHMT consensus sequence. Most of the conserved glycines are located at predicted turns, indicating a probable tertiary structural conservation (5). This suggests that these glycines are structurally important, being located at turns, at positions where helices cross, or at other constrained locations. The importance of having a glycine at five of these positions is proven by this work. In allele *t1576*, glycine 372 is changed to arginine; in allele *t1679*, glycine 143 is changed to aspartic acid; in alleles *t1552*, *t1632*, and *t1665*, conserved glycines 204, 406, and 149, respectively, are all changed to glutamic acids. Replacing these constrained glycines with large charged groups probably has a deleterious effect on the structure of SHMT.

Two of the mutations cause alanine to valine alterations, *t1666* (alanine 103) and *s2518* (alanine 126). These residues are conserved alanines in the SHMT consensus sequence. We assume there are very strong structural constraints at these positions, as the conservative change of an alanine to a valine disrupts the structure enough to affect SHMT activity. A single PCR product could not be obtained for allele *t1666* with primer set B in the heteroduplex analysis, but the point mutation does not interfere with the primer binding site (Fig. 3.). To eliminate the possibility that *t1666* contained a second mutation in one of the B primer binding sites, the overlapping regions, A and C, were sequenced. Neither region contained a mutation, so the failure of this allele to yield a single PCR product with the B primer set remains a mystery.

Allele *t1473* contains a double mutation, as suggested by the heteroduplex analysis where region A and B both indicated the position of a mutation. Alanine 63 is changed to valine, and leucine 146 is changed to phenylalanine. Residue 63 is an alanine four out of six times in the SHMT consensus sequence, but the *E. coli* protein has an arginine at this position, so we assume that changing this alanine to valine would have a minimal steric effect. However, residue 146 is a conserved leucine in the SHMT consensus sequence, suggesting that the leucine to phenylalanine missense mutation is the primary cause of the mutant phenotype in allele *t1473*.

In alleles *t1671* and *t1520*, serine 251 is changed to phenylalanine. Residue 251 is only three amino acids away from the active site and is always small in the SHMT consensus sequence. The replacement of this small amino acid with a large aromatic one probably distorts the conformation of the active site and may also interfere with the PLP aromatic ring.

In allele *t1607*, histidine 259 is changed to tyrosine. The mutation in allele *t1607*, residue 259 in the SHMT consensus sequence, is the only one from our collection that has been mutated previously. This histidine is in the conserved active site VTTTTHK(S/T) motif found in all SHMTs and is adjacent to the active site lysine. Every residue in this motif, except the valine, has been mutated in the *E. coli* isoform of SHMT to determine the effects on catalysis (Fig. 1.). When the active site histidine is changed to asparagine, there is no structural change in the enzyme, but the catalytic activity is greatly

SHMT). The alignment was made using the ClustalW program. In the SHMT consensus sequence (*consensus*), *uppercase letters* indicate amino acids identical in all six proteins, *lowercase letters* indicate residues identical in at least four out of six proteins, and a *dot* indicates conserved amino acids. The *numbers* indicate the positions in the *C. elegans* protein. The amino acids mutated in the *mel-32* alleles are in *bold*, and the mutant amino acid for each is listed above the sequence. Residues that have been previously mutated in *E. coli* are *boxed* (see text for details).

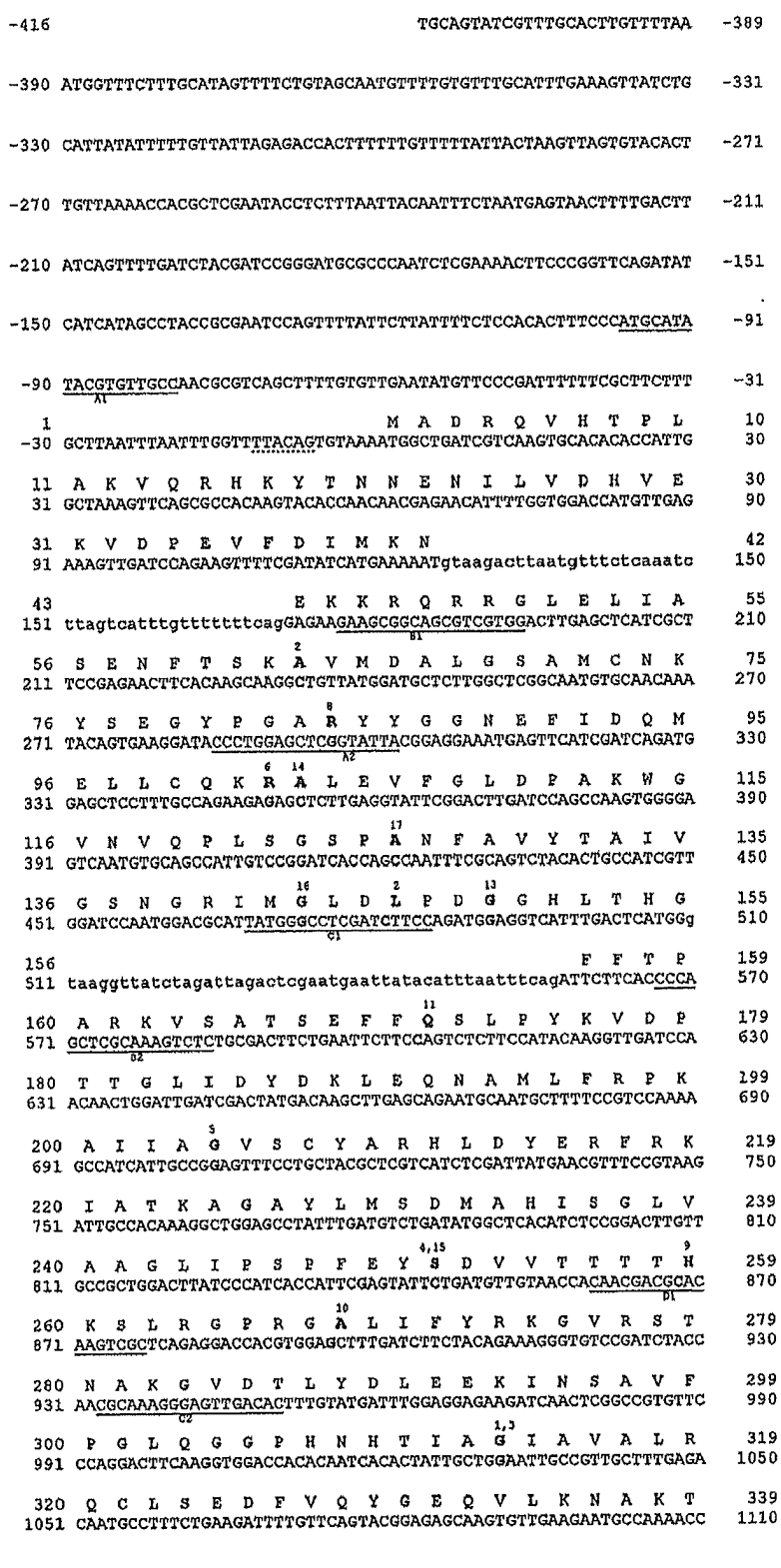


FIG. 3. **Nucleotide sequence of the *C. elegans* SHMT genomic region and derived amino acid sequence.** Intronic regions are in *lowercase*, the primers used in the heteroduplex and sequence analysis are *underlined*, a putative SL1 leader sequence is denoted by a *dashed line*, and the poly(A) signal is *double underlined*. The nucleotides and amino acids mutated in the *mel-32* alleles are in *bold*, and the allele number for each is listed. This sequence corresponds to positions 39507–37321 (in reverse orientation) of the C05D11 GenBank entry (accession number U00048).

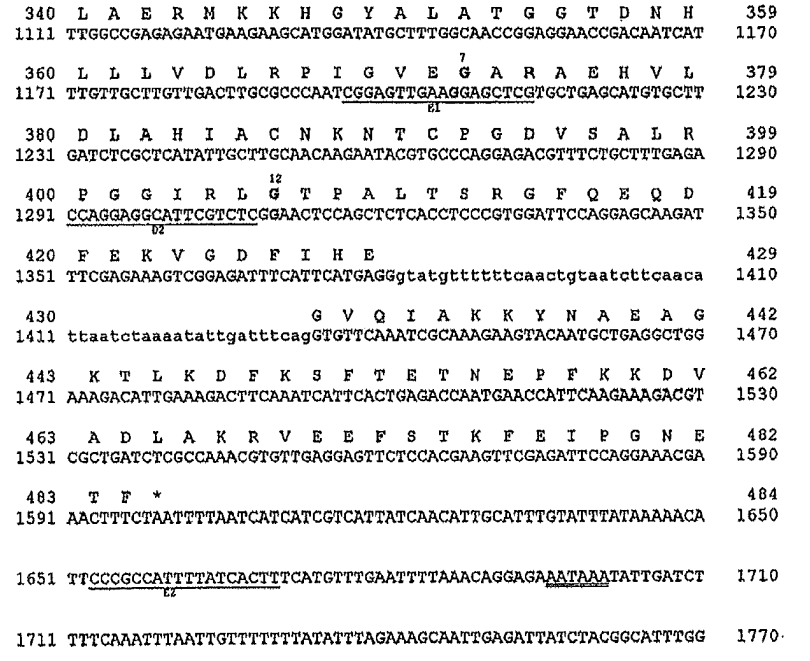


FIG. 3—*continued*

reduced (25). When the active site histidine is changed to aspartic acid, the activity of the physiological reactions is reduced, but the activity of some alternate reactions is increased (26). A series of kinetic and spectral studies on these mutant enzymes revealed that the active site histidine is not catalytically essential and is not the base that accepts the α proton. These studies have also shown that the active site histidine interacts with the amino acid substrate or PLP. This histidine is believed to have a critical role in determining reaction specificity by determining the structure of the one-carbon binding site and controlling the orientation of the substrate and PLP ring (25, 26). Because *E. coli* SHMT is catalytically active without an active site histidine, it is unlikely that the imidazole ring of histidine 259 participates as an electron donor or acceptor in the physiological reactions of CeSHMT. In allele *t1607*, inserting the aromatic ring of tyrosine at the active site may change the environment of the PLP aromatic ring such that it is no longer in line with critical residues required for catalysis. The PLP and substrate binding pocket would also be distorted.

The *E. coli* active site is VTTTTHKT. Each of the threonine residues in this motif was changed to an alanine to determine the kinetic and spectral properties of the mutant enzymes (27). When the first or fourth threonine is converted to alanine, the enzyme is essentially wild-type. When the second or fifth threonine is changed to alanine, the mutants are structurally unchanged but have shifted kinetic properties. These results indicated that these four threonines do not play a critical role in the mechanism of SHMT. However, when the third threonine is mutated to an alanine, the enzyme loses 97% of its catalytic activity. When this residue is mutated to a serine, the activity is essentially wild-type, so the presence of a hydroxyl-containing side chain is very important. These studies also revealed that none of these threonines bind PLP (27).

The active site lysine in *E. coli* has also been mutated. When this lysine is changed to a glutamine, the enzyme catalyzes one turnover of product at wild-type levels, but cannot release the product (28). When the lysine is changed to arginine or histidine, the PLP cofactor cannot readily form the external aldimine. These results suggest that the active site lysine expels the product by converting the external aldimine to an internal aldimine and that lysine is not the base that removes the α proton of the substrate.

It has been proposed that there are two bases at the active site of SHMT on opposite sides of the PLP ring (29). These bases are not the active site histidine (25, 26, 29) or lysine (28). A study on sheep liver SHMT has shown that there are arginine residues present at the active site. Arginine 269 and arginine 462 from sheep liver were protected from chemical modification by tetrahydrofolate binding (33). Arginine 462 is not conserved but arginine 269 (residue 273 in CeSHMT) is conserved as an arginine in all eukaryotic SHMTs and as a lysine in *E. coli* SHMT. Arginine 363 and arginine 372 in *E. coli* SHMT (conserved residues 404 and 413 in CeSHMT, Fig. 1.) were changed to both alanine and lysine (30). Both of the arginine 372 mutations had wild-type activity, suggesting that this residue, although conserved, is not critical for catalytic activity. The R363A mutant enzyme had no activity with serine as a substrate and could not bind serine or glycine. The R363K mutant enzyme had only 0.03% of wild-type activity and a 15-fold decreased affinity for serine and glycine. The conserved arginine at this position is catalytically essential and believed to be the binding site of the amino acid substrate carboxyl group (30). It has been proposed that an arginine-carboxyl interaction might be preferred over simple charge interactions because the guanidium group presents charged hydrogen bonds rather than the single bond formed by lysine or histidine (33). This may explain why replacing an arginine with a lysine has such a dramatic effect on enzyme activity.

The mutant phenotype of allele *t1555* is caused by an arginine (Arg-102) to lysine mutation. The arginine at this position is 100% conserved in all sequenced SHMTs. This conservation suggests that the presence of the guanidine group of arginine at this position is essential to the enzymatic function of SHMT. In allele *t1597*, arginine 84 is changed to glutamine. This arginine is also 100% conserved, so the presence of a basic group at this position is probably also essential to the enzymatic function of

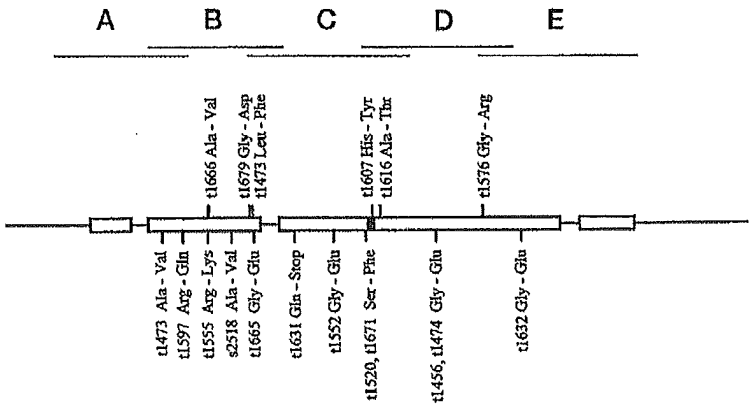


FIG. 4. **Exon/intron structure of *mel-32*(SHMT) and summary of sequencing results.** The overlapping PCR products used for heteroduplex analysis are shown above the gene. The wild-type and mutant amino acid(s) are given for each allele. The active site is indicated by a *shaded box*.

TABLE II
mel-32 mutant alleles

	Allele	Codon mutation	Amino acid mutation
1	*t1456*	GGA → GAA	Gly-313 → Glu
2	*t1473*	GCT → GTT	Ala-68 → Val
		CTT → TTT	Leu-146 → Phe
3	*t1474*	GGA → GAA	Gly-313 → Glu
4	*t1520*	TCT → TTT	Ser-251 → Phe
5	*t1552*	GGA → GAA	Gly-204 → Glu
6	*t1555*	AGA → AAA	Arg-102 → Lys
7	*t1576*	GGA → AGA	Gly-372 → Arg
8	*t1597*	CGG → CAG	Arg-84 → Gln
9	*t1607*	CAC → TAC	His-259 → Tyr
10	*t1616*	GCT → ACT	Ala-268 → Thr
11	*t1631*	CAG → TAG	Gln-171 → Stop
12	*t1632*	GGA → GAA	Gly-406 → Glu
13	*t1665*	GGA → GAA	Gly-149 → Glu
14	*t1666*	GCT → GTT	Ala-103 → Val
15	*t1671*	TCT → TTT	Ser-251 → Phe
16	*t1679*	GGC → GAC	Gly-143 → Asp
17	*s2518*	GCC → GTC	Ala-126 → Val

SHMT. Either or both of these arginine residues could be at the active site of the enzyme.

Chemical modification studies on the cysteine residues in rabbit liver cytosolic (31) and mitochondrial (32) SHMT, sheep liver cytosolic SHMT (5), and *E. coli* SHMT (34) have shown that there are no disulfide bonds in eukaryotic homotetramers or prokaryotic homodimers, and that PLP and substrates protect a catalytically essential active site cysteine. There are no cysteine mutations in our collection, but residue 99 in the SHMT consensus sequence is a conserved cysteine in all the eukaryotic isoforms. This residue could be the important active site cysteine.

In a series of experiments to study folding intermediates of *E. coli* SHMT, a set of mutant proteins was constructed in which the three tryptophan residues were replaced with phenylalanine residues (35). The three double mutants and a triple mutant were essentially wild-type enzymes, with only minor structural differences. Tryptophans 16, 183, and 385 in *E. coli* correspond to positions 37, 214, and 426 in the SHMT consensus sequence (Fig. 1). These tryptophans are not conserved, but, intriguingly, position 37 in the SHMT consensus sequence contains five out of six aromatic residues; position 214 has six out of six aromatic residues, with all five eukaryotic forms having a tyrosine; position 426 also has six out of six aromatic residues, tryptophan in *E. coli* and a mixture of phenylalanine and tyrosine in the eukaryotic forms. Clearly, the presence of an aromatic side chain at these positions is important.

It is assumed that each of the *mel-32* mutations abolishes, or greatly reduces, the SHMT activity, resulting in the observed maternal effect lethal phenotype. We hypothesize that a product of SHMT, such as glycine or some byproduct of the one-carbon metabolic pool, is required for embryonic development in the egg. The essential metabolite is normally supplied by the diet but cannot diffuse through the eggshell. *mel-32* homozygotes have enough maternally provided SHMT activity to develop and hatch into feeding larvae, where the diet can supply enough of the required nutrient. These *mel-32* homozygotes, however, have no SHMT activity to pass on to their offspring, which quickly use up the required metabolite pool in the egg and arrest as embryos. This hypothesis can also explain why *C. elegans* is unique in its requirement for SHMT activity for survival. The SHMT mutations in *E. coli*, *S. cerevisiae*, and Chinese hamster ovary cells all result in glycine auxotrophy (3, 4, 9, 10). In all three cases, the deficiency is rescued by addition of glycine. The egg shell of *C. elegans*, which is impervious to most chemicals (36), would imprison the embryo in a forced starvation. This hypothesis can be tested, and the missing metabolite identified, by supplying the developing larvae with metabolic precursors, for example glycine and thymidine.

The 17 point mutations found in *C. elegans mel-32*(SHMT) are clustered within the middle two thirds of the protein, the most conserved region. This suggests that the ends of the protein are not required for enzymatic function and that any point mutations occurring there would not be detected in a screen for maternal effect lethals, the likely null phenotype. This observation is supported by the fact that if the amino-terminal 25–30 amino acids of rabbit liver cytosolic SHMT are removed with proteases the enzyme remains catalytically active and structurally stable (37). Sixteen of the *mel-32*(SHMT) alleles contain a single base pair substitution, and one, *t1473*, contains two base pair substitutions. All of the mutations are G/C to A/T transitions, as expected in EMS-induced alleles.

mel-32 has 17 alleles, making it a high frequency hit gene, as most Mel genes only have one or two alleles. Several large scale screens for maternal effect lethals have been carried out which have produced hundreds of Mel alleles (38–40).[6,7] Most of these mutations have only been identified genetically, but a few have been cloned and sequenced. Most of the identified Mel genes are involved in polarizing the embryo or determining cell fate. For example, *par* (for partitioning defective) mutant embryos arrest as amorphous masses of differentiated cells (39). *par-1* encodes a conserved Ser/Thr kinase, and the two alleles that have been sequenced reveal mutations in invariant kinase domain residues, suggesting that PAR-1 kinase activity has an essential function (41). *par-2* encodes a 628-amino acid protein with a putative ATP binding site and zinc ring finger domain. Two sequenced alleles of *par-2* introduce stop codons that trun-

cate the protein to a form lacking the ATP binding domain (42). The *pie-1* gene encodes a zinc finger protein, which interferes with transcription (43). The only sequenced allele of this gene contains a 217-bp deletion in the 5′ end of the gene (44). The *mex-1* gene contains two copies of the zinc finger domain found in *pie-1* and mutations in *mex-1* alter the fates of some somatic blastomeres. Two deletion mutants have been sequenced; one deletes the NH_2-terminal 36 amino acids, and the other deletes the COOH-terminal 80 amino acids (45). The *mex-3* gene contains two RNA binding KH domains and regulates blastomere identity in embryos. One allele of *mex-3* is deleted for the first 92 bp of coding sequence, whereas three other alleles have point mutations that change conserved glycines in the KH domain (46). The transcription factor *skn-1* is required for correct specification of cell fates. One mutant form of *skn-1* has a stop codon introduced in the DNA binding domain (47, 48). Finally, the *apx-1* gene of *C. elegans* controls early cell fates and is a homolog of the *Drosophila Delta* and *Serrate* genes. One mutant allele of *apx-1* has a 900-bp deletion in the center of the gene (49).

The extreme degree of conservation in SHMT across millions of years of evolution suggests that the conserved residues must play some catalytically or structurally essential roles. Several of these important residues have previously been identified. From our collection of 17 mutations, only one, the active site histidine, has been mutated previously. Therefore, this work defines 13 new residues (not including the nonsense and two repeated mutations) that are potentially essential to SHMT's catalytic activity or structural integrity.

SHMT is believed to have a folding pattern similar to aspartate aminotransferase and dialkylglycine decarboxylate, both of which have crystallographic structure data available (50). Three out of four of the residues conserved in all aminotransferases are conserved in SHMT. In the CeSHMT numbering scheme, these residues are: aspartic acid 231, which hydrogen bonds to N1 of PLP; lysine 260, which forms a Schiff base with PLP; and arginine 404, which hydrogen bonds with the α-carboxyl group of the substrate. Future crystallographic analysis of SHMT may allow characterization of more essential residues. The residues corresponding to arginine 404 and lysine 260 have been mutated in *E. coli* (Fig. 1) and shown to be essential (28, 30), but there are no known mutations affecting aspartic acid 231.

Injecting SHMT subclones containing site-directed mutations for aspartic acid 231, and other conserved residues, into *mel-32*(SHMT) nulls may allow a quick assay of enzymatic function. If the mutated subclone rescues, then the mutant SHMT is functional; if it does not rescue, then we can assume the mutation has abolished or decreased SHMT function. *C. elegans* could also be used to test chemotherapy drugs directed against SHMT. If shells of wild-type eggs are solubilized, and the embryos treated with drugs, any drug that reduces or eliminates SHMT function should cause a Mel phenotype.

Acknowledgments—We thank Diana Janke, Jacquie Schein, and The Ha for expert technical assistance and Raja Rosenbluth for advice.

REFERENCES

1. Schirch, L. (1982) *Adv. Enzymol.* **53**, 83–112
2. Garrow, T. A., Brenner, A. A., Whitehead, V. M., Chen, X. N., Duncan, R. G., Korenberg, J. R., and Shane, B. (1993) *J. Biol. Chem.* **268**, 11910–11916
3. Narkewicz, M. R., Sauls, S. D., Tjoa, S. S., Teng, C., and Fennessey, P. V. (1996) *Biochem. J.* **313**, 991–996
4. McNeil, J. B., McIntosh, E. M., Taylor, B. V., Zhang, R., Tang, S., and Bognar, A. L. (1994) *J. Biol. Chem.* **269**, 9155–9165
5. Usha, R., Savithri, H. S., and Rao, N. A. (1994) *Biochim. Biophys. Acta* **1204**, 75–83
6. Eichler, H. G., Hubbard, R., and Snell, K. (1981) *Biosci. Rep.* **1**, 101–106
7. Nakshatri, H., Bouillet, P., Bhat-Nakshatri, P., and Chambon, P. (1996) *Gene (Amst.)* **174**, 79–84
8. Trakatellis, A., Dimitriadou, A., Exindari, M., Christodoulou, D., Malissiovas, N., Antoniadis, A., and Haitoglou K. (1994) *Postgrad. Med. J.* **70**, (Suppl. 1), S89–S92
9. Stauffer, G. V., Plamann, M. D., and Stauffer, L. T. (1981) *Gene (Amst.)* **14**, 63–72
10. Pfender, W., and Pizer, L. I. (1980) *Arch. Biochem. Biophys.* **200**, 503–512
11. Stover, P. J., Chen, L. H., Suh, J. R., Stover, D. M., Keyomarsi, K., and Shane, B. (1997) *J. Biol. Chem.* **272**, 1842–1848
12. Brenner, S. (1974) *Genetics* **77**, 71–94
13. Durbin, R., and Thierry-Mieg, J. (1997) *C. elegans II*, p. 17, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
14. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci U. S. A.* **85**, 2444–2448
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410
16. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
17. Rychlik, W., and Rhoads, R. E. (1989) *Nucleic Acids Res.* **17**, 8543–8551
18. Kramer, J. M., French, R. P., Park, E.-C., and Johnson, J. J. (1990) *Mol. Cell. Biol.* **10**, 2081–2089
19. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) *EMBO J.* **10**, 3959–3970
20. Barstead, R. J., Kleiman, L., and Waterston, R. H. (1991) *Cell Motil. Cytoskel.* **20**, 69–78
21. Williams, B. D., Schrank, B., Huynh, C., Shownkeen, R., and Waterston, R. H. (1992) *Genetics* **131**, 609–624
22. Nijbroek, G., Sood, S., McIntosh, I., Francomano, C. A., Bull, E., Pereira, L., Ramirez, F., Pyeritz, R. E., and Dietz, H. C. (1995) *Am. J. Hum. Genet.* **57**, 8–21
23. Janke, D. L., Schein, J. E., Ha, T., Franz, N. W., O'Neil, N. J., Vatcher, G. P., Stewart, H., I., Kuervers, L. K., Baillie, D. L., and Rose, A. M. (1997) *Genome Res.* **7**, 974–985
24. Keen, J., Lester, D., Inglehearn, C., Curtis, A., and Bhattacharya, S. (1991) *Trends Genet.* **7**, 5
25. Hopkins, S., and Schirch, V. (1986) *J. Biol. Chem.* **261**, 3363–3369
26. Stover, P., Zamora, M., Shostak, K., Gautam-Basak, M., and Schirch, V. (1992) *J. Biol. Chem.* **267**, 17679–17687
27. Angelaccio, S., Pascarella, S., Fattori, E., Bossa, F., Strong, W., and Schirch, V. (1992) *Biochemistry* **31**, 155–162
28. Schirch, D., Delle Fratte, S., Iurescia, S., Angelaccio, S., Contestabile, R., Bossa, F., and Schirch, V. (1993) *J. Biol. Chem.* **268**, 23132–23138
29. Shostak, K., and Schirch, V. (1988) *Biochemistry* **27**, 8007–8014
30. Delle Fratte, S., Iurescia, S., Angelaccio, S., Bossa, F., and Schirch, V. (1994) *Eur. J. Biochem.* **225**, 395–401
31. Gavilanes, F., Peterson, D., and Schirch, L. (1982) *J. Biol. Chem.* **257**, 11431–11436
32. Gavilanes, F., Peterson, D., Bullis, B., and Schirch, L. (1983) *J. Biol. Chem.* **258**, 13155–13159
33. Usha, R. Savithri, H. S., and Rao, N. A. (1992) *J. Biol. Chem.* **267**, 9289–9293
34. Schirch, V., Hopkins, S., Villar, E., and Angelaccio, S. (1985) *J. Bacteriol.* **163**, 1–7
35. Cai, K., and Schirch, V. (1996) *J. Biol. Chem.* **271**, 2987–2994
36. Wood, W. B. (1988) *The Nematode Caenorhabditis elegans*, p. 217, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
37. Schirch, V., Schirch, D., Martini, F., and Bossa, F. (1986) *Eur. J. Biochem.* **161**, 45–49
38. Kemphues, K. J., Kusch, M., and Wolf, N. (1988) *Genetics* **120**, 977–986
39. Kemphues, K. J., Priess, J., Morton, D., and Cheng, N. (1988) *Cell* **52**, 311–320
40. Mains, P. E., Sulston, I. A., and Wood, W. B. (1990) *Genetics* **125**, 351–369
41. Guo, S., and Kemphues, K. J. (1995) *Cell* **81**, 611–620
42. Levitan, D. J., Boyd, L., Mello, C. C., Kemphues, K. J., and Stinchcomb, D. T. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6108–6112
43. Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., Priess, J. R., and Fire, A. (1996) *Nature* **382**, 713–716
44. Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J. R. (1996) *Nature* **382**, 710–712
45. Guedes, S., and Priess, J. R. (1997) *Development* **124**, 731–739
46. Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J., and Priess, J. R. (1996) *Cell* **87**, 205–216
47. Bowerman, B., Eaton, B. A., and Priess, J. R. (1992) *Cell* **68**, 1061–1075
48. Blackwell, T. K., Bowerman, B., Priess, J. R., and Weintraub, H. (1994) *Science* **266**, 621–628
49. Mello, C. C., Draper, B. W., and Priess, J. R. (1994) *Cell* **77**, 95–106
50. Pascarella, S., Schirch, V., and Bossa, F. (1993) *FEBS Lett.* **331**, 145–149

Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase

Younes ACHOURI*, Mark H. RIDER†, Emile VAN SCHAFTINGEN* and Mariette ROBBI*‡

*Laboratory of Physiological Chemistry and †Hormone and Metabolic Research Unit, International Institute of Cellular and Molecular Pathology and Université Catholique de Louvain, B-1200 Brussels, Belgium

Rat liver D-3-phosphoglycerate dehydrogenase was purified to homogeneity and digested with trypsin, and the sequences of two peptides were determined. This sequence information was used to screen a rat hepatoma cDNA library. Among 11 positive clones, two covered the whole coding sequence. The deduced amino acid sequence (533 residues; M_r 56493) shared closer similarity with *Bacillus subtilis* 3-phosphoglycerate dehydrogenase than with the enzymes from *Escherichia coli*, *Haemophilus influenzae* and *Saccharomyces cerevisiae*. In all cases the similarity was most apparent in the substrate- and NAD^+-binding domains, and low or insignificant in the C-terminal domain. A corresponding 2.1 kb mRNA was present in rat tissues including kidney, brain and testis, whatever the dietary status, and also in livers of animals fed a protein-free, carbohydrate-rich diet, but not in livers of control rats, suggesting transcriptional regulation. The full-length rat 3-phosphoglycerate dehydrogenase was expressed in *E. coli* and purified. The recombinant enzyme and the protein purified from liver displayed hyperbolic kinetics with respect to 3-phosphoglycerate, NAD^+ and NADH, but substrate inhibition by 3-phosphohydroxypyruvate was observed; this inhibition was antagonized by salts. Similar properties were observed with a truncated form of 3-phosphoglycerate dehydrogenase lacking the C-terminal domain, indicating that the latter is not implicated in substrate inhibition or in salt effects. By contrast with the bacterial enzyme, rat 3-phosphoglycerate dehydrogenase did not catalyse the reduction of 2-oxoglutarate, indicating that this enzyme is not involved in human D- or L-hydroxyglutaric aciduria.

INTRODUCTION

D-3-Phosphoglycerate dehydrogenase (EC 1.1.1.95) catalyses the first step in the pathway of serine formation from glycolytic intermediates (reviewed in [1]). This enzyme is widely distributed in organisms and in tissues. In rat liver, its activity depends strongly on nutritional status, being low in animals fed a normal diet and increasing more than 10-fold upon ingestion of a low-protein, carbohydrate-rich diet [2,3]. When measured in the non-physiological direction, the enzyme present in human fibroblasts is inhibited by concentrations of the substrate 3-phosphohydroxypyruvate above 10 μM, and this inhibition is released by salts [4]. Similar properties have been described for bovine liver D-glycerate dehydrogenase, the inhibitory substrate being hydroxypyruvate in this case [5,6], but were not observed with 3-phosphoglycerate dehydrogenases from bacteria [7] or plants [8]. In *Escherichia coli* [9], *Bacillus subtilis* [7] and plants [8], 3-phosphoglycerate dehydrogenase is subject to allosteric control by the terminal product of the pathway, serine, whereas such control does not appear to be present in the animal enzyme [10].

Recently, the possible involvement of 3-phosphoglycerate dehydrogenase in the pathogenesis of two newly described hereditary disorders has attracted our attention. The activity of this enzyme was indeed markedly decreased (to 13 and 22% of the normal value) in fibroblasts from two patients with a decreased concentration of serine in the plasma and cerebrospinal fluid [4]. Furthermore, *E. coli* 3-phosphoglycerate dehydrogenase was shown to catalyse the reduction of 2-oxoglutarate to both the D- and L-isomers of 2-hydroxyglutarate, suggesting that a mutation of the human enzyme may contribute to the neurometabolic diseases D- and L-hydroxyglutaric aciduria [11].

The primary sequences of 3-phosphoglycerate dehydrogenases from *E. coli* [12], *Haemophilus influenzae* (GenBank accession no. L45106; [13]), *Saccharomyces cerevisiae* (GenBank P40054) and *Bacillus subtilis* (GenBank L47648; [14]) are known, as is the three-dimensional structure of the *E. coli* enzyme [15]. Each subunit of the tetrameric protein has three distinct domains: a nucleotide-binding domain (residues 108–294), a substrate-binding domain (residues 7–107 and 295–336) and a regulatory domain, which binds L-serine (residues 337–410). The main contact points between the subunits are at the level of the coenzyme-binding domains and the regulatory domains, indicating the importance of these zones for the tetramerization of the enzyme.

The aim of the present work was to initiate an investigation of the mammalian enzyme at the molecular level, and to study its kinetics and specificity, particularly its ability to catalyse the reduction of 2-oxoglutarate to 2-hydroxyglutarate, and the mechanism by which its activity is controlled by diet.

MATERIALS AND METHODS

Materials

3-Phosphoglycerate, NAD^+, NADH, the random prime DNA labelling kit and *Taq* and *Pwo* DNA polymerases were from Boehringer Mannheim. 3-Phosphohydroxypyruvate was prepared from the tricyclohexylammonium salt of the dimethyl ketal derivative [16] purchased from Sigma. [α-^{32}P]dCTP (approx. 3000 Ci/mmol), [α-^{35}S]dATP (> 1000 Ci/mmol), the Δ Taq Cycle Sequencing kit (USB) and Hybond N® membranes were from Amersham International. The T7 Sequencing® kit was

‡ To whom correspondence should be addressed.

The sequence of the *R. norvegicus* cDNA encoding 3-phosphoglycerate dehydrogenase has been deposited in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession no. X97772 RND3PGDEH.

from Pharmacia Biotech Inc., and the IRD41 dye-labelled M13 reverse primer was from LI-COR.

Enzyme purification from rat liver and amino acid sequencing of two tryptic peptides

Male Wistar rats were fed on cornflour for 2 weeks before being killed to increase the expression of the liver enzyme [2]. The livers of 14 rats were homogenized in 4 vol. of 0.25 M sucrose containing 20 mM Hepes, pH 7.5, 100 mM KCl, 1 mM dithiothreitol and 0.25 mM PMSF. The homogenate was spun for 30 min at 9000 *g* and 4 °C. A 7–22 % (w/v) poly(ethylene glycol) 8000 fraction was prepared from this extract and dissolved in 200 ml of 10 mM Hepes, pH 7.5, 1 mM dithiothreitol and 1 mM EDTA (buffer A) supplemented with 0.25 mM PMSF. The preparation was clarified by centrifugation and loaded on to a DEAE-Sepharose column (2.5 cm × 22 cm; flow rate 2.5 ml/min). The column was washed with 100 ml of buffer A containing 0.25 mM PMSF and eluted with a linear salt gradient (0–400 mM KCl in 400 ml of buffer A containing 0.25 mM PMSF, 1 μg/ml leupeptin and 1 μg/ml antipain). The enzyme was then diluted 13-fold in buffer A and loaded on to an AMP–Sepharose column (1 cm × 3 cm) equilibrated with buffer A. The enzyme was eluted from the affinity column with 0.125 mM NAD^+ and 0.125 mM 3-phosphoglycerate in buffer A. In order to obtain internal tryptic peptides, the band corresponding to 3-phosphoglycerate dehydrogenase in the purified preparation was cut from a Coomassie Blue-stained SDS/polyacrylamide gel and the protein was concentrated in agarose for trypsin digestion [17]. Peptides were purified by narrow-bore HPLC for microsequencing [17].

Enzyme assays

3-Phosphoglycerate dehydrogenase was assayed spectrophotometrically at 30 °C in a mixture comprising 25 mM Hepes, pH 7.1, 90 μM phosphohydroxypyruvate, 90 μM NADH and 400 mM KCl, unless otherwise indicated. One unit is the amount of enzyme catalysing the reduction of 1 μmol of phosphohydroxypyruvate/min under these assay conditions. For the determination of kinetic constants in the physiological direction, activity was measured in a mixture containing 200 mM Tris, pH 9, 180 mM hydrazine sulphate, 1 mM EDTA and the indicated concentrations of 3-phosphoglycerate and NAD^+.

Amplification of a cDNA fragment by PCR with degenerate oligonucleotides

A library constructed from rat FTO2B hepatoma cells in Uni-Zap® XR (Stratagene), with oligo(dT)-primed cDNAs inserted between *Eco*RI and *Xho*I restriction sites, was kindly provided by V. J. Dupriez (I.C.P. Brussels) [18]. The DNA purified from this library was used as a template to amplify a cDNA fragment with *Taq* DNA polymerase. The sense primer [5′-GC(C/T)G-G(A/C)AC(A/C)GG(A/C)GT(G/C)GA(C/T)AA(C/T)GT-(G/C)GA-3′] corresponded to the first nine amino acids (AGTGVDNVD) of one of the sequenced peptides (see Figure 1); the antisense primer was a 20-mer corresponding to the vector-specific T7 promoter. The amplified product (1450 bp) was purified, cloned in a T vector prepared from pBluescript [19] and sequenced.

cDNA cloning and sequencing

About 300000 plaques from the cDNA library described above were plated and lifted on to Hybond N® membranes. The DNA was cross-linked to the dried filters by UV irradiation. Prehybrization for 1 h and hybridization for 15 h were performed at 65 °C in a solution containing 6× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 5× Denhardt's solution [20], 0.5 % SDS and 20 μg/ml denatured herring sperm DNA. An approx. 1000 bp *Xho*I restriction fragment corresponding to the 3′ end of the cDNA was radiolabelled with [α-^{32}P]dCTP by random priming [20] and used as a probe. The membranes were washed at 65 °C (2× SSC, 3 × 10 min; 2× SSC/0.1 % SDS, 1 × 10 min; 2× SSC, 1 × 10 min; 0.2× SSC, 1 × 10 min), dried and exposed to Kodak XAR-5 film at room temperature for 2–18 h.

Recombinant pBluescript cDNAs were excised *in vivo* from the lambda Uni-ZAP® XR vector as recommended by the manufacturer and purified by the alkaline lysis method [20]. The plasmid containing the longest insert was sequenced completely in both directions by the dideoxy method [21] with the T7 Sequencing® kit, [α-^{35}S]dATP and specific primers from within the cDNA sequence and from flanking regions of the vector. In some cases, the Δ Taq Cycle Sequencing kit® and the IR-dye-labelled M13 reverse primer were used and the products were analysed using an automated laser fluorescence DNA sequencer 4000L from LI-COR. Sequences were analysed using the DNA Strider programme [22]. Multiple sequence alignment was performed using the GenBank program PILEUP.

Northern blots

Total RNA was isolated from various tissues by the guanidinium isothiocyanate/CsCl procedure, subjected to electrophoresis in 1 % agarose/formaldehyde gels and transferred by capillarity to nylon membranes [20]. These were hybridized and washed as described above.

Expression and purification of recombinant 3-phosphoglycerate dehydrogenase

The coding region of the full-length cDNA clone of 3-phosphoglycerate dehydrogenase was amplified by PCR using *Pwo* DNA polymerase, a low-error enzyme, with one primer (GGAATTC-CATATG GCCTTCGCAAATCTG) containing the start codon (underlined) inserted in a *Nde*I site, and a second primer (CGGGATCCTTCAGAAGCAGAACTGGAA) containing the stop codon (underlined) flanked by a *Bam*HI site. The amplified fragment was cloned into pBluescript restricted with *Eco*RV, excised from this vector with *Nde*I and *Bam*HI and inserted into the expression vector pET3a [23] to produce the recombinant plasmid pET3PGDH. For the preparation of the shortened form of the enzyme, the second primer used in the amplification (GGTGTGGGATCCTCAGGCACTGGTAAG) was designed to replace Phe-325 by a stop codon (underlined). To produce the recombinant enzymes, BL21(DE3)p*Lys*S cells transformed with the expression plasmids were aerobically grown in 1 litre of LB medium containing 0.2 % glucose and 50 μg/ml ampicillin at 37 °C up to an A_{600} of 0.6 unit. The culture was then quenched for 30 min on ice before addition of 0.4 mM isopropyl β-D-thiogalactoside to induce expression of the cloned enzyme. Culture was continued for another 18 h at 18 °C. Cells were harvested by centrifugation (12000 *g*, 30 min, 4 °C) and extracted as described in [24]. The recombinant enzyme was purified by chromatography on DEAE-Sepharose and AMP–Sepharose, as described for the native liver enzyme. The active fractions were diluted with 1 vol. of glycerol and stored at −20 °C, under which conditions the enzyme was stable for several months.

Protein was determined by the procedure of Bradford [25], with bovine γ-globulin as a standard.

RESULTS AND DISCUSSION

Purification of rat liver 3-phosphoglycerate dehydrogenase and screening of the cDNA library

Rat liver 3-phosphoglycerate dehydrogenase was purified to homogeneity from the livers of rats that had been fed for 2 weeks with a protein-free, high-carbohydrate diet, by a procedure derived from that used by Lund et al. [26]. The purified preparation had a specific activity of 35 units/mg of protein at 30 °C and contained a single polypeptide chain of M_r 57000 in SDS/polyacrylamide gels (results not shown). The protein was concentrated and digested with trypsin. Tryptic peptides were purified, two of which gave the following sequences: AGTGVDNVDLEAATR (peptide 1) and ALESGECAGAALDVFTEEPP (peptide 2). PCR-amplification with a primer derived from peptide 1, a primer corresponding to the T7 promoter and DNA from a hepatoma cDNA library yielded an approx. 1450 bp fragment, which was cloned and found by sequencing to encode a protein homologous to *E. coli* 3-phosphoglycerate dehydrogenase.

This ~ 1450 bp fragment was used to screen the hepatoma cDNA library. The 11 positive clones that were obtained had inserts of between 0.63 and 1.78 kb that were differently truncated at their 5′ ends. The largest clone was sequenced over its entire length on both strands.

Nucleotide sequence and deduced amino acid sequence

The nucleotide and predicted amino acid sequences are shown in Figure 1. The proposed ATG start codon lies within a suitable consensus sequence for translation initiation by eukaryotic ribosomes [27]. It is preceded by an in-frame TAA stop codon and opens a reading frame encoding a protein of 533 amino acids with a predicted M_r of 56493 and a pI of 6.26. These values are in agreement with the behaviour of the liver enzyme in SDS/PAGE (57000 M_r) and in chromatography on ion exchangers. The enzyme is indeed retained by DEAE-Sepharose at pH 7.5 and by SP-Sepharose at pH 6.0 (results not shown). The next ATG codon is situated about 285 bp downstream. Its use would result in a peptide with a much lower M_r than observed (approx. 45000). The open reading frame is preceded by 68 bp of 5′-untranslated region and followed by a 3′-untranslated sequence of 113 bp containing an 18 bp poly(A) tail. The putative, somewhat atypical, poly(A) addition signal AGTAAA starts 18 bp upstream of the poly(A) tail.

The amino acid sequences of the two peptides resulting from digestion with trypsin (Figure 1) match the sequence of the recombinant enzyme, with the exception of two glutamine residues that were detected as glutamic acid during sequencing of the second peptide, presumably as a result of deamidation. These data indicate that the cloned cDNA encodes the subunit of the liver D-3-phosphoglycerate dehydrogenase. Partial sequences obtained from the other, incomplete, cDNA clones were consistent with that shown in Figure 1, indicating that the enzyme is expressed as a single molecular form in FTO2B cells.

Comparison with related enzymes from other species

The predicted sequence of the rat liver enzyme has been aligned with 3-phosphoglycerate dehydrogenases from *E. coli*, *H. influenzae*, *S. cerevisiae* and *B. subtilis* [12–14] (Figure 2). The Figure indicates the positions of different domains in the *E. coli* enzyme, as well as amino acids implicated in catalysis. GenBank also contains the sequences of mouse and human partial cDNAs encoding peptides displaying > 90% identity with rat liver 3-phosphoglycerate dehydrogenase (not shown).

```
GCCTTCAGTTTCCTGTACTAAGTGCTTCTGCCCACCAGAGCAACCGATTCTAAGGCCTGG      60
CTCTAGCAATGGCCTTCGCAAATCTGCGCAAAATACTCATCAGTGATAGCCTCGACCCCT     120
          M  A  F  A  N  L  R  K  I  L  I  S  D  S  L  D  P   17
GCTGCCGGAAGATCCTGCAAGATGGAGGGCTGCAGGTGGTGGAGAAGCAGAACTTGAGCA     180
C  C  R  K  I  L  Q  D  G  G  L  Q  V  V  E  K  Q  N  L  S   37
AGGAGGAGCTGATAGCCGAACTCCAGGACTGTGAAGGCCTTATCGTCCGGTCAGCTACTA     240
K  E  E  L  I  A  E  L  Q  D  C  E  G  L  I  V  R  S  A  T   57
AGGTCACTGCTGATGTCATCAATGCAGCAGAGAAGCTCCAGGTGGTGGGCAGGGCTGGTA     300
K  V  T  A  D  V  I  N  A  A  E  K  L  Q  V  V  G  R  A  G   77
CAGGCGTGGACAATGTGGATCTGGAGGCTGCCACAAGGAAGGGCGTCCTCGTCATGAACA     360
T  G  V  D  N  V  D  L  E  A  A  T  R  K  G  V  L  V  M  N   97
CCCCCAATGGAAATAGCCTCAGTGCTGCGGAACTCACCTGTGGGATGCTCATGTGCCTCG     420
T  P  N  G  N  S  L  S  A  E  L  T  C  G  M  L  M  C  L  L  117
CCAGGCAGATCCCCCAGGCGACGGCTTCGATGAAAGATGGCAAATGGGACCGGAAGAAGT     480
A  R  Q  I  P  Q  A  T  A  S  M  K  D  G  K  W  D  R  K  K  137
TCATGGGGACAGAGCTGAACGGGAAGACACTGGGAATTCTTGGCCTGGGCAGAATTGGAA     540
F  M  G  T  E  L  N  G  K  T  L  G  I  L  G  L  G  R  I  G  157
GAGAGGTGGCCGCCCGAATGCAGGCCTTTGGAATGAAGACTGTAGGCTATGACCCCATCA     600
R  E  V  A  A  R  M  Q  A  F  G  M  K  T  V  G  Y  D  P  I  177
TTTCTCCAGAAGTCGCTGCCTCCTTTGGTGTTCAGCAGCTGCCGCTAGAGGAGATCTGGC     660
I  S  P  E  V  A  A  S  F  G  V  Q  Q  L  P  L  E  E  I  W  197
CTCTCTGTGATTTCATCACTGTCCATACCCCGCTCCTGCCCTCCACTACAGGCTTGCTCA     720
P  L  C  D  F  I  T  V  H  T  P  L  L  P  S  T  T  G  L  L  217
ATGACAGCACCTTTGCCCAGTGCAAGAAAGGCGTGCGGGTGGTGAACTGTGCTCGAGGAG     780
N  D  S  T  F  A  Q  C  K  K  G  V  R  V  V  N  C  A  R  G  237
GCATTGTGGATGAAGGTGCCCTGCTCCGTGCCCTGCAGTCTGGTCAGTGTGCTGGTGCTG     840
G  I  V  D  E  G  A  L  L  R  A  L  Q  S  G  Q  C  A  G  A  257
CACTGGATGTGTTTACAGAAGAGCCACCACGGGACCGGGCCTTAGTGGACCACGAGAACG     900
A  L  D  V  F  T  E  E  P  P  R  D  R  A  L  V  D  H  E  N  277
TCATCAGCTGTCCCCACCTGGGCGCCAGCACCAAGGAGGCCCAGAGCCGCTGTGGGGAGG     960
V  I  S  C  P  H  L  G  A  S  T  K  E  A  Q  S  R  C  G  E  297
AAATCGCAGTCCAGTTTGTGGACATGGTGAAGGGGAAATCTCTAACAGGGGTTGTAAACG    1020
E  I  A  V  Q  F  V  D  M  V  K  G  K  S  L  T  G  V  V  N  317
CCCAGGCTCTTACCAGTGCCTTCTCTCCACACACCAAGCCTTGGATTGGTCTGGCAGAAG    1080
A  Q  A  L  T  S  A  F  S  P  H  T  K  P  W  I  G  L  A  E  337
CATTGGGCACGCTGATGCACGCCTGGGCTGGCCTCCCCTAAAGGGACCATCCAGGTGGTGA    1140
A  L  G  T  L  M  H  A  W  A  G  S  P  X  G  T  I  Q  V  V  357
CACAAGGAACATCTCTGAAGAATGCTGGGACCTGCCTGAGCCCTGCGGTCATTGTCGCGCC    1200
T  Q  G  T  S  L  K  N  A  G  T  C  L  S  P  A  V  I  V  G  377
TTCTGAGAGAAGCATCAAAACAGGCAGATGTGAACTTGGTGAACGCTAAGCTACTGGTGA    1260
L  L  R  E  A  S  K  Q  A  D  V  N  L  V  N  A  K  L  L  V  397
AAGAGGCTGGCCTCAATGTCACCACCTCCCACAGTCCTGGTGTCCCAGGAGAGCAGGGCA    1320
K  E  A  G  L  N  V  T  T  S  H  S  P  G  V  P  G  E  Q  G  417
TCGGGGAATGCCTCCTGACTGTGGCCTTGGCAGGTGCCCCCTACCAAGCTGTGGGCTTGG    1380
I  G  E  C  L  L  T  V  A  L  A  G  A  P  Y  Q  A  V  G  L  437
TCCAGGGCACCACACCAATGTTGCAGATGCTCAACGGAGCTGTCTTCAGGCCAGAGGTGC    1440
V  Q  G  T  T  P  M  L  Q  M  L  N  G  A  V  F  R  P  E  V  457
CTCTACGCAGGGGCCAGCCCCTGCTCCTGTTCCGGGCTCAGCCCTCCGACCCTGTCATGC    1500
P  L  R  R  G  Q  P  L  L  L  P  R  A  Q  P  S  D  P  V  M  477
TGCCCACTATGATCGGCCTACTGGCAGAGGCGGGGGTACAGCTGCTGTCCTACCAGACCT    1560
L  P  T  M  I  G  L  I  A  E  A  G  V  Q  L  L  S  Y  Q  T  497
CCAAGGTGTCTGACGGAGACACTTGGCACGTCATGGGCCTCTCCTCCCTACTGCCCAGCC    1620
S  K  V  S  D  G  D  T  W  H  V  M  G  L  S  S  L  L  P  S  517
TGGACGCATGGAAGCAGCATGTTTCTGAGGCTTTCCAGTTCTGCTTCTGACCCAGGGGCT    1680
L  D  A  W  K  Q  H  V  S  E  A  F  Q  F  C  F            533
CAGCGGTCCCAGCCCCTCAGGCTCTTCTGAGGAAACCCGCTCACTGTGACCTGAACTAAT    1740
ATCTAGTAAAGAATCTAACTCCAAAAAAAAAAAAAAAAA                        1780
```

Figure 1 Sequence of the cDNA encoding rat hepatoma 3-phosphoglycerate dehydrogenase and its deduced amino acid sequence

The initiator ATG codon, the stop codon and the putative polyadenylation signal are underlined. The in-frame TAA stop codon upstream of the Met initiator codon is shown in italics. The amino acid sequences corresponding to the two tryptic peptides are underlined. Note that the two predicted glutamine residues of the second peptide were recovered as glutamate by amino acid sequencing.

Remarkably, the rat enzyme shows closer similarity to its *B. subtilis* counterpart than to the three other enzymes. The rat and *B. subtilis* 3-phosphoglycerate dehydrogenases are of about the same length, which is about 120 residues longer than either the *E. coli* or *H. influenzae* enzymes. This difference is mainly due to the presence of a longer C-terminal domain (~ 200 residues, as compared with ~ 75 residues). Furthermore, the nucleotide- and substrate-binding domains of the rat enzyme show more identity with the *B. subtilis* enzyme (46%) than with the other enzymes (32%). For the *E. coli*/*S. cerevisiae* pair, the corresponding value is even higher (53%), indicating that the yeast enzyme is closer to the *E. coli* enzyme than to the mammalian enzyme. The percentage identity is much lower in the regulatory domain than in the other domains, and is 15% for the rat/*B. subtilis* pair, 30% for the *E. coli*/*S. cerevisiae* pair and not significant when the rat enzyme is compared with that from *E. coli*. These observations suggest that there are two different types of 3-phosphoglycerate dehydrogenase that originated from a duplication event occurring before the separation of eukaryotes and prokaryotes.

Figure 2 also shows that the consensus sequence Gly-Xaa-Gly-Xaa-Xaa-Gly-Xaa$_{17}$-Asp (Figure 2), involved in binding the

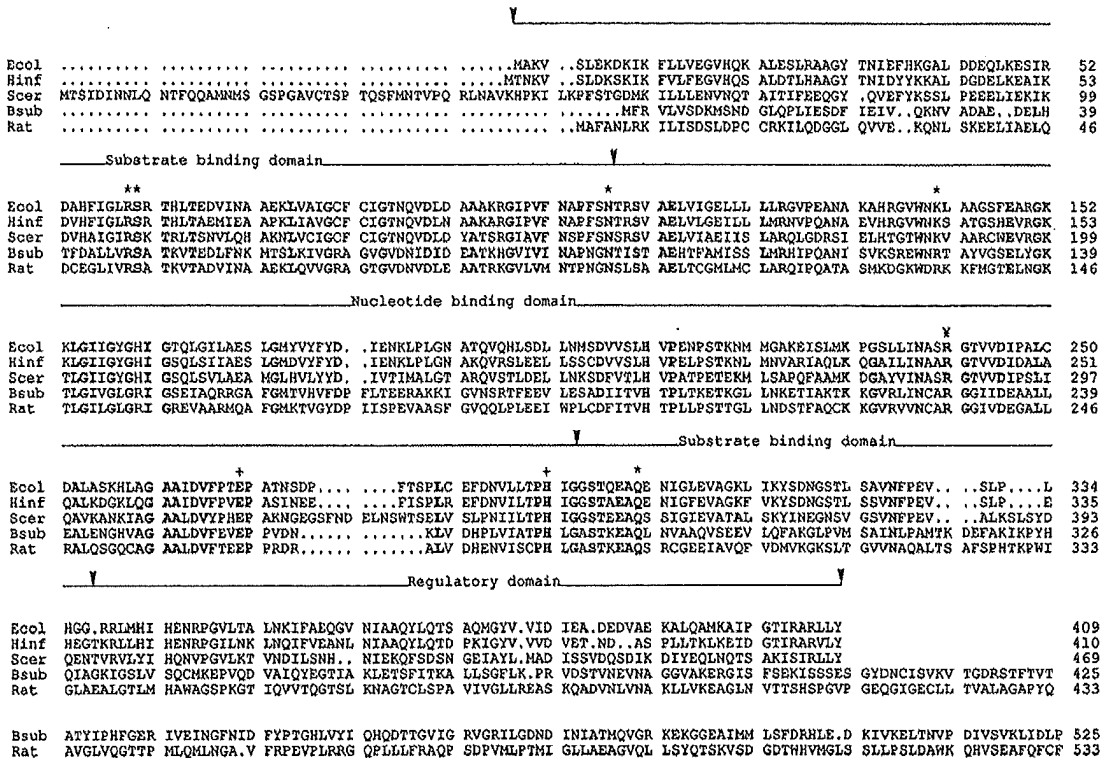


Figure 2 Alignment of the sequences of 3-phosphoglycerate dehydrogenase from various species

The amino acid sequences of the enzyme from *E. coli* (Ecol; [12]), *H. influenzae* (Hinf; GenBank accession no. L45106; [13]), *S. cerevisiae* (Scer; GenBank P40054), *B. subtilis* (Bsub; GenBank L47648; [14]) and rat hepatoma (the present work) are displayed in single-letter code after alignment for maximal identity. Gaps are indicated by dots and strictly conserved amino acids are in bold. Asterisks indicate residues surrounding the catalytic pocket, where hydride transfer occurs. Residues proposed to be involved in proton shuttling (+) and the arginine presumably involved in binding of the carboxylate group of the substrate (Y) [15] are also shown. The domains of the *E. coli* enzyme are shown above the alignment.

adenosine portion of NAD^+ [28], is located in the rat enzyme at residues 152–175. Several residues lining the catalytic pocket in the *E. coli* enzyme are strictly conserved in the mammalian enzyme (Arg-60, Ser-61, Asn-108 and Gln-301) or replaced by functionally similar residues (Lys-141 replaced by Arg-136). By analogy with the *E. coli* enzyme, it is likely that Arg-236 binds the carboxy group of the substrate and that His-292 and Glu-269 function as a proton shuttle [15].

Northern blots

Northern blot analysis of total RNA indicated the presence of a specific ~ 2.1 kb mRNA in livers derived from animals maintained for 3 days on a protein-free, high-carbohydrate diet (Figure 3). The length of this mRNA is consistent with the size of the longest clone that we obtained, allowing for the presence of the poly(A) tail. Strikingly, the mRNA was not detectable in the livers of animals fed on a normal diet, suggesting that control of the expression of 3-phosphoglycerate dehydrogenase is exerted at the level of transcription or by stabilization of the mRNA. A 2.1 kb mRNA was also observed in kidney (Figure 3), testis and brain (results not shown), and in lower amounts in lung (Figure 3), skeletal muscle and heart (results not shown). In these tissues, no effect of dietary manipulation on 3-phosphoglycerate dehydrogenase mRNA was observed. These results are in agreement with the observation that the induction of the enzyme by a low-protein diet is restricted to the liver [29].

Expression of recombinant rat 3-phosphoglycerate deydrogenase

The isolated cDNA was expressed in *E. coli* to verify that it encoded a protein with similar properties to the enzyme purified from liver. Addition of isopropyl thiogalactoside to BL21-(DE)p*Lys*S cells carrying the expression plasmid resulted in the induction of 3-phosphoglycerate dehydrogenase, the activity of which reached about 8 units/mg of protein after 18 h at 18 °C (as compared with 0.01 unit/mg of protein in control cells). The recombinant protein was purified by chromatography on DEAE-Sepharose, and was nearly homogeneous after this single step, corresponding to a polypeptide with the expected size (M_r 57000 by SDS/PAGE; results not shown). After AMP–Sepharose chromatography, the preparation had a specific activity of 35 units/mg of protein. From 1 litre of culture, 4700 units of purified enzyme was obtained, corresponding to a 50% yield. Upon gel filtration on Sephacryl S-200, the enzyme was eluted with an M_r of approx. 220000 (results not shown), indicating that it behaves as a tetramer, like the enzyme from *E. coli* [30], rabbit

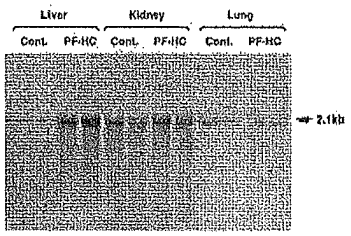


Figure 3 Northern blot analysis of mRNA from liver, kidney and lung from rats fed either a normal diet or a protein-free, carbohydrate-rich diet

Rats either were fed laboratory pellets (Cont.) or were maintained for 3 days on a protein-free, high-carbohydrate diet (PF-HC). Total RNA was extracted and subjected to electrophoresis (15 μg/lane). For each condition, samples from two different animals are shown.

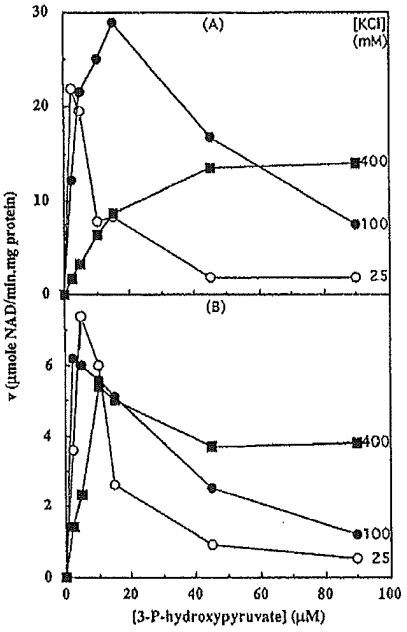


Figure 4 Effects of 3-phosphohydroxypyruvate and KCl on the activity of the recombinant full-length (A) and truncated (B) forms of liver 3-phosphoglycerate dehydrogenase

The enzyme was assayed with the indicated concentrations of phosphohydroxypyruvate (3-P-hydroxypyruvate) and KCl, in the presence of 90 μM NADH.

liver [26] and chicken liver [31]. The recombinant protein purified on DEAE-Sepharose was used in the kinetic studies.

Kinetic properties

The kinetic properties of the recombinant and native enzymes were investigated. When measured in the physiological direction, both enzymes showed K_m values of about 100 μM for 3-phosphoglycerate and 27 μM for NAD^+. In the opposite direction, the enzyme displayed a K_m of 20–25 μM for NADH, and was inhibited by elevated concentrations of 3-phosphohydroxypyruvate (Figure 4A). As previously reported by Jaeken et al. [4], this inhibition was released by 100–400 mM KCl. The data obtained in the present study at low substrate concentrations indicate that the salt displaced the saturation curve to the right, inhibiting enzyme activity at low concentrations of substrate and stimulating activity at elevated concentrations. Similar effects were obtained with other salts. When measured with 90 μM 3-phosphohydroxypyruvate, the effect was biphasic, with an optimum at a concentration close to 400 mM for univalent salts and 80 mM for bivalent salts (results not shown).

We have also tested the effects of serine, which is known to inhibit the activity of the 3-phosphoglycerate dehydrogenases from *E. coli* [9], *B. subtilis* [7] and plants [8], and of the other 19 standard amino acids at concentrations up to 5 mM, on the activity both in the physiological direction and in the reverse direction; none of these amino acids affected the enzymic activity. Since the enzyme from rat appears to be closer to the *B. subtilis* enzyme than to the other enzymes, we took the precaution of testing the sensitivity to serine under conditions similar to those described by Saski and Pizer [7], i.e. by preincubating the enzyme with serine in the presence of dithiothreitol. No effect was observed under these conditions.

Properties of a truncated form of rat 3-phosphoglycerate dehydrogenase

Because the inhibition by excess substrate appeared to be specific for the mammalian enzyme, we tested the hypothesis that the C-terminal domain mediates this inhibition. We therefore expressed in *E. coli* a truncated form of 3-phosphoglycerate dehydrogenase lacking the last 209 amino acids. The recombinant protein was purified to a specific activity of 14 units/mg of protein. As expected, its subunit migrated with an M_r of 36000 in SDS/PAGE (results not shown). Gel filtration on Sephacryl S-200 gave an M_r of 80000, indicating that the enzyme behaved as a dimer and confirming the role of the C-terminal domain in the tetramerization of the enzyme [15]. However, the mutant enzyme displayed substrate inhibition by phosphohydroxypyruvate, and this inhibition was sensitive to salt (Figure 4B). These results indicate that the C-terminal domain of the rat enzyme is not involved in the inhibition by excess phosphohydroxypyruvate.

Lack of 2-oxoglutarate reductase activity

Since *E. coli* 3-phosphoglycerate dehydrogenase has been reported to catalyse the reduction of 2-oxoglutarate to L- and D-hydroxyglutarate [11], we tested whether the rat enzyme could catalyse a similar reaction. However, at concentrations between 10 μM and 10 mM, 2-oxoglutarate was not a substrate for rat 3-phosphoglycerate dehydrogenase, whether 400 mM NaCl was present or not. Furthermore, at 10 mM, 2-oxoglutarate did not behave as an inhibitor of the enzyme in either the forward or reverse direction. Similar results were observed with the truncated enzyme. These results therefore argue against the possibility that 3-phosphoglycerate dehydrogenase is implicated in the pathogenesis of D- or L-hydroxyglutaric aciduria [11]. The results indicate also that 3-phosphoglycerate dehydrogenase is not responsible for the L-2-hydroxyglutarate dehydrogenase activity found in rat liver [32].

Conclusion

This paper describes the cloning of the cDNA encoding rat liver 3-phosphoglycerate dehydrogenase and the confirmation of its identity by expression of the encoded protein in bacteria. The availability of this clone now opens up the possibility of studying the molecular basis of human 3-phosphoglycerate dehydrogenase

deficiency and the mechanism by which protein intake regulates the expression of this enzyme in the liver.

We thank H. G. Hers for a critical reading of the manuscript, J. F. Collet for his help in database searching and H. Degand (Laboratory of Physiological Biochemistry, University of Louvain) for performing the microsequencing. This work was supported by the Actions de Recherche Concertées and by the Belgian Federal Service for Scientific, Technical and Cultural Affairs. M.H.R. is a Chercheur Qualifié of the Belgian Fonds National de la Recherche Scientifique.

REFERENCES

1 Snell, K. (1986) Trends Biochem. Sci. **11**, 241–243
2 Fallon, H. J., Hackney, E. J. and Byrne, W. L. (1966) J. Biol. Chem. **241**, 4157–4167
3 Mauron, J., Mottu, F. and Spohr, G. (1973) Eur. J. Biochem. **32**, 331–342
4 Jaeken, J., Detheux, M., Van Maldergem, L., Foulon, M., Carchon, H. and Van Schaftingen, E. (1996) Arch. Dis. Child. **74**, 542–545
5 Sugimoto, E. and Pizer, L. I. (1968) J. Biol. Chem. **243**, 2081–2089
6 Coderch, R., Lluis, C. and Bozal, J. (1979) Biochim. Biophys. Acta **566**, 21–31
7 Saski, R. and Pizer, L. I. (1975) Eur. J. Biochem. **51**, 415–427
8 Slaughter, J. C. and Davies, D. D. (1968) Biochem. J. **109**, 749–755
9 Pizer, L. I. (1963) J. Biol. Chem. **238**, 3934–3944
10 Walsh, D. A. and Sallach, H. J. (1965) Biochemistry **4**, 1076–1085
11 Zhao, G. and Winkler, M. E. (1996) J. Bacteriol. **178**, 232–239
12 Tobey, K. L. and Grant, G. A. (1986) J. Biol. Chem. **261**, 12179–12183
13 Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M. et al. (1995) Science **269**, 496–512
14 Sorokin, A., Zumstein, E., Azevedo, V., Ehrlich, S. D. and Serror, P. (1993) Mol. Microbiol. **10**, 385–395
15 Schuller, D. J., Grant, G. A. and Banaszak, L. J. (1995) Nature Struct. Biol. **2**, 69–76
16 Ballou, C. E. (1960) Biochem. Prep. **7**, 66–68
17 Rider, M. H., Puype, M., Van Damme, J., Gevaert, K., De Boeck, S., D'Alayer, J., Rasmussen, H. H., Celis, J. E. and Vandekerckhove, J. (1995) Eur. J. Biochem. **230**, 258–265
18 Dupriez, V. J., Darville, M. I., Antoine, I. V., Gégonne, A., Ghysdael, J. and Rousseau, G. G. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 8224–8228
19 Marchuk, D., Drumm, M., Saulino, A. and Collins, F. S. (1991) Nucleic Acids Res. **19**, 1154
20 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
22 Marck, C. (1988) Nucleic Acids Res. **16**, 1829–1836
23 Studier, F. W. and Moffatt, B. A. (1986) J. Mol. Biol. **189**, 113–130
24 Veiga-da-Cunha, M., Detheux, M., Watelet, N. and Van Schaftingen, E. (1994) Eur. J. Biochem. **225**, 43–51
25 Bradford, M. M. (1976) Anal. Biochem. **72**, 248–254
26 Lund, K., Merrill, D. K. and Guynn, R. W. (1986) Biochem. J. **238**, 919–922
27 Kozak, M. (1986) Cell **44**, 283–292
28 Wierenga, R. K., Terpstra, P. and Hol, W. G. J. (1986) J. Mol. Biol. **187**, 101–107
29 Hayashi, S., Tanaka, T., Naito, J. and Suda, M. (1975) J. Biochem. (Tokyo) **77**, 207–219
30 Rosenbloom, J., Sugimoto, E. and Pizer, L. I. (1968) J. Biol. Chem. **243**, 2099–2107
31 Grant, G. A. and Bradshaw, R. A. (1978) J. Biol. Chem. **253**, 2727–2731
32 Jansen, G. A. and Wanders, R. J. (1993) Biochim. Biophys. Acta **1225**, 53–56

Received 2 October 1996; accepted 3 December 1996